Department of Electrical and
Computer Engineering

Villanova University

FINAL REPORT

(11/15/97 - 9/30/01)

BLIND TIME-FREQUENCY ANALYSIS FOR SOURCE DISCRIMINATION IN MULTISENSOR ARRAY PROCESSING

Submitted to

Office of Naval Research

Grant No. N00014-98-1-0176

Principal Investigator

Moeness G. Amin

Contributors

Prof. Moeness Amin
Dr. Yimin Zhang
Dr. Adel Belouchrani
Mr. Weifeng Mu

October 2001

Table of Contents

BLIND TIME-FREQUENCY ANALYSIS FOR SOURCE DISCRIMINATION IN MULTISENSOR ARRAY PROCESSING

1. EXECUTIVE SUMMARY

This report presents a comprehensive summary of the research work performed under the ONR funding, grant number N00014-98-1-0176 over the period of November 15th, 1997 to September 30th, 2001. The research team over periods of the grant consists of Prof. Moeness Amin (PI), Dr. Adel Belouchrani (Postdoctoral Fellow), Dr. Yimin Zhang (Postdoctoral Fellow), and Mr. Weifeng Mu (Graduate Student).

The results obtained in this project have been disseminated over the past four years in nineteen journal articles and over twenty-five peer reviewed conference papers. Section 2 lists, with complete citations, all publications acknowledging the ONR support. This list is also used to point the reader of the executive summary to the pertinent papers of each sub-section. As most of the research has already been documented in three annual reports (October 1998, October 1999, October 2000), we include in section 4 of this final report detailed descriptions of the main contributions over the past four years highlighted in nine key journal publications. Section 3 includes an introduction to the theory underlying the thrust of our approach to solving array signal processing problems using time-frequency distributions.

In addition to the research team at Villanova University, we have collaborated with scientists from Canada, Turkey, Spain, and Australia on several important aspects and objectives of the ONR project.

1.1 Main Contributions

The research fund from the ONR over the past four years has been effectively used for advancing the theory of time-varying spectrum analysis and its applications to the important area of sensor array processing. We pioneered the development of multi-sensor receivers based on quadratic time-frequency and joint-variable distributions. We have provided the theoretical framework for solving direction finding and blind source separation problems using bilinear transforms, and established the fundamental role of time-frequency distributions for the rejection of nonstationary interferers in direct sequence spread spectrum communication systems. We were the first to utilize the time-frequency signatures of signal arrivals for improved detection, angle-of-arrival estimation, and blind signal recovery. We were also the first to use the signal power distributions in the time-frequency domain for optimum excision of smart jammers and a large class of undesired waveforms through time-varying filtering, least-square synthesis methods, and subspace projection techniques. The research team working on ONR funding at Villanova University have introduced the original concepts of spatial time-frequency distributions, and more generally spatial joint-variable distributions, and provided the comprehensive analysis of their role in high resolution angle-of-arrival estimation and equalization of the communication channel. In doing so, we have elevated bilinear transformations from a purely temporal processing tool to a more powerful spatio-temporal processing technique. In this respect, we opened the door for time-frequency distributions to join cyclostationary, second- and higher-order statistics methods in utilizing the spatial dimension for enhanced receiver performance. Our work has shown that time-frequency distributions represent the natural mechanism to process narrowband signals in broadband communication platforms, stationary signals in a dominating nonstationary environment, and spread spectrum communication signals contaminated by powerful jammers.

1.2 Fast Computational Time-Frequency Distribution Kernels

Time-frequency distributions (TFDs) are a powerful tool for nonstationary signal analysis in a wide variety of applications. However, the computational burden limits the applicability of the TFD in practice. Hence, design of fast implementation methods are important to utilize TFDs for applications that require on-line processing such as radar systems and nonstationary jammer excision in spread spectrum communications. The TFD is often computed using the following procedure. First, the bilinear data products are computed which form the instantaneous autocorrelation function (IAF). Next, the IAF is smoothed along each time-lag using a time-frequency (t-f) kernel to produce the local autocorrelation function (LAF). Finally, the LAF is Fourier transformed along the time-lag variable to yield the required TFD. Recognizing the fact that the computations required to compute the DFT are much smaller than those needed to compute the LAF, it becomes logical to target the calculation of the LAF, for significant computational savings.

We have introduced a new recursive structure for efficient computations of the local autocorrelation function. This recursion is achieved by decomposing the t-f kernel along each time-lag into a finite number of cosinusoidal terms. Each term generates one recursive equation to update the LAF. This technique is named the *trigonometric decomposition of time-frequency kernels*, and can be viewed as a frequency sampling design technique of t-f kernels. This proposed class of Trigonometric kernels (also referred to as the frequency sampling (FS) kernels), while providing considerable reduction in the computation of the LAF, retain the Hermition property of the LAF which simplifies the DFT operations. Further, trigonometric decomposition is amenable to data -dependent kernel design and does not compromise the time-frequency desirable properties.

Publications[36,38,40]

1.3 Blind Source Separation in the Ambiguity Domain

Time-frequency distributions are proposed for applications to array signal processing problems. For this purpose, spatial time-frequency distributions (STFDs) have been introduced and represented in a matrix form. The elements of a spatial time-frequency distribution matrix are the time-frequency distributions and the cross-time frequency distributions of the data received at the multi-sensor array. It has been shown that the relationship between the spatial time-frequency distributions of the sensor data and the time-frequency distributions of the sources is identical to that of the sensor data covariance matrix and the sources' correlation matrix. This key property permits direction finding and blind source separations to be performed using the sources' time-frequency localization properties.

Blind source separation consists of recovering a set of signals of which only instantaneous linear mixtures are observed. This problem has been typically solved using statistical information available on source signals, including second or higher order statistics. Our primary contribution in this area is to show that the spatial time-frequency distributions is an effective alternative to separating sources whose signatures are different in the t-f domain. Successful applications of STFDs to source separation require computing STFD at different time-frequency (t-f) points. The results are then incorporated into a joint-diagnalization technique to estimate the mixing, or the array manifold matrix. This matrix transfers the source signals into sensor data – a process that we must *undo* by blind separation, often up to a multiplicative complex scalar and the order of the sources.

Next, we have improved source signal estimation by performing the blind source separation using ambiguity functions rather than time-frequency distributions. Unlike the spatial time-frequency distributions, where the separation is performed using various time-frequency bins, the proposed approach is based on information from the ambiguity domain. The procedure is similar to that adopted in spatial TFD-based blind source separation. The main difference, however, lies in the domain in which the mixing matrix is estimated. In the proposed technique, the mixing matrix is obtained by joint diagonalization of the spatial ambiguity (SAF) matrices rather than the spatial time-frequency distribution matrices. These SAF matrices are formed from the auto- and cross- ambiguity functions of the data received by the different array sensors. It is shown that at each time-lag and frequency-lag point in the ambiguity domain, the sensors' SAF matrix is related to the sources' ambiguity function matrix in the same fashion as the sensors' and sources' TFDs. Performing blind source separation in the ambiguity-domain, rather than the t-f domain, provides a greater ability to formulate the problem using the signal auto-terms. Avoiding cross-terms both simplifies and improves the performance of joint-diagonalization, and subsequently enhances source estimation.

Publications[18,21,26,29,33,37,39,42]

1.4 Spatial Averaging of Time-Frequency Distribution for Source Separation

Symmetric spatial averaging of spatial time-frequency distributions has been introduced. The spatial averaging of the spatial time-frequency distributions of the data across an antenna array removes the undesired effect of crossterms between the impinging signals. These terms reside along the off-diagonal entries of the source time-frequency distribution matrix, and consequently impede the source separation performance, which is based on pre-assumed diagonal matrix structure. Spatial averaging amounts to forming a spatial Hermitian Toeplitze matrix using the auto- and cross-time-frequency distributions of the data over one half of the array. This matrix is then added to the spatial matrix corresponding to the other half of the array. The desired effect of this averaging is reallocating the interaction between the source signals in the time-frequency domain from the off diagonal to the diagonal elements of the source TFD matrix. In this respect, cross-terms, due to their potential high values, are regarded as useful components that could be employed for improved performance. Without spatial-averaging, array performance is very sensitive to whether only auto-term or cross-term points or their mix are incorporated in the source separation procedure. With spatial averaging, this is no longer a concern, and as such, a major burden is using bilinear time-frequency distributions has been alleviated.

Publications[13,31]

1.5 Bilinear Signal Synthesis in Antenna Arrays

We have introduced a new approach for signal synthesis in antenna arrays using the spatial separation of the sources as well as the sources' localization properties in the time-frequency domain. In effect, source separation, or signal recovery, is performed based on the difference in both the time-frequency and spatial signatures of the signal arrivals. In the proposed approach, we permit the source spatial signatures to play a fundamental role in enhancing the true signal t-f power concentration, leading to improved synthesis performance. This is achieved by averaging of the time-frequency distributions of the sensor data across the array prior to proceeding with commonly used least squares bilinear syntheses methods. The signals impinging

on the receiver are assumed to be localizable in the time-frequency domain, e.g., FM and polynomial phase signals.

Unlike the existing source separation techniques based on time-frequency signal representations, the devised approach does not require whitening or retrieval of the source directional matrix, thereby, simplifies the signal recovery process. Further, as a result of the averaging process across the array, a weighting function in the time-frequency domain is constructed which decreases the noise levels, reduces the interactions of the source signals, and mitigates the cross-terms. These desirable effects are obtained independent of the temporal characteristics of the source signals and without causing any smearing of the signal auto-terms.

Publications[4,6,7,14]

1.6 Time-Frequency MUSIC for Direction Finding

Our first contribution to direction finding is the introduction the Time-frequency MUSIC as a new array signal processing method based on time-frequency signal representations. As discussed above, spatial time-frequency distributions have successfully been used to solve the problem of blind source separations for nonstationary signals. We have used the same underlying structure of STFDs to solve the direction finding problem, i.e., angles of arrival (AOA) estimation. A new method for the estimation of the signal subspace and the noise subspace based on time-frequency signal representations has been introduced. The proposed approach consists of the joint block-diagonalization (JBD) of a set of spatial time-frequency distribution matrices. Once the signal and the noise subspaces are estimated, any subspace based approach can be applied for AOA estimation. We propose to use the MUSIC algorithm. The effects of spreading the noise power while localizing the source energy in the time-frequency domain amounts to increasing the robustness of the eigenstructure superresolution method with respect to noise, and hence improves spatial resolution. Performance of the proposed Time-Frequency MUSIC (TF-MUSIC) for different time-frequency kernels was evaluated numerically.

Publications[12,24,28,32,35,43]

1.7 Time-Frequency Maximum Likelihood Methods for Direction Finding

We have introduced a novel time-frequency maximum likelihood (t-f ML) method for direction-of-arrival (DOA) estimation for non-stationary signals impinging on a multi-sensor array receiver. We have shown the superiority of this method over conventional maximum likelihood DOA estimation techniques. Time-frequency distributions localize the signal power in the time-frequency domain, and as such enhance the effective SNR, leading to improved DOA estimation. The localization of signals with different time-frequency signatures permits the division of the time-frequency domain into smaller regions; each contains smaller number of signals than those incident on the array. The reduction of the number of signals within different time-frequency regions not only reduces the required number of sensors, but also decreases the computational load in multi-dimensional optimizations. Compared to the recently proposed time-frequency MUSIC (t-f MUSIC), the proposed t-f ML method can be applied in coherent environments, without the need to perform any type of preprocessing that is subject to both array geometry and array aperture.

Publications[17,27]

1.8 A Wideband Perspective to Spatial Time-Frequency Distributions

All previous work on this project has relied on the assumption of narrowband arrays. That is, the spatial signature of a nonstatinoary signal, e.g., a chirp signal, is assumed constant over the source bandwidth. This may be true for cases in which the signal bandwidth is negligible compared to the carrier frequency. Although simplifies the analysis and simulations, the narrowband assumption may not always hold true, specifically, in military signal processing applications.

We have successfully relaxed the narrowband array constraint, imposed in past algorithm developments, and extended the narrowband spatial time-frequency distributions (STFDs) to the wideband case. A new STFD-based wideband root-MUSIC estimator is proposed. This technique employs an extended coherent signal-subspace (CSS) principle involving coherent averaging over a pre-selected set of time-frequency points rather than the conventional frequency-only averaging procedure. For FM signals, it is demonstrated that the proposed method outperforms both the conventional, commonly used, CSS algorithm for wideband sources and is also superior to the iterative root-MUSIC technique recently appeared in the literature.

Publications[19,20]

1.9 Subspace Analysis of Spatial Time-Frequency Distribution Matrices

Subspace analysis of spatial time-frequency distribution (STFD) matrices have been developed. It has been shown that for signals with clearly defined time-frequency signatures, such as FM signals, smaller estimation errors in the signal and noise subspaces can be achieved by using spatial time-frequency matrices over the subspace estimates obtained from using the data covariance matrix approach. This improvement in subspace estimation is the result of incorporating the time-frequency points along the instantaneous frequencies of the impinging signals on the array into the subspace estimation procedure. These points belong to autoterm regions of high power concentrations, and as such, when used in constructing STFDs, they provide high SNR matrices with improved eigen-decompositions. The advantages of STFD-based direction finding over traditional direction finding methods using data covariance matrices were demonstrated using the MUSIC algorithm. It was shown that the time-frequency MUSIC outperforms conventional MUSIC in the two situations of low SNR and closely spaced sources. Unlike conventional array processing techniques, which are nondiscriminatory, and must therefore spatially localize all signals incident on the array, the STFD-based array processing provides the flexibility of dealing with all signal arrivals, or a subset of them. In this respect, it does not suffer from the drawback of requiring higher number of sensors than sources. The ability to select fewer sources depends on the distinction of their time-frequency signatures from those of other source signals. The eigenstructure of the STFD matrix constructed from the time-frequency points that belong to the autoterm regions of a number of sources will only yield the signal subspace of these sources. It was shown that the maximum improvement in subspace estimation using STFD over data covariance matrices is achieved when constructing the STFD from only one source signal.

Publications[9]

1.10 Chirp Beamforming

The problem of direction finding in sensor arrays in the presence of constant modulus signals is addressed. A new deterministic maximum likelihood (ML) direction of arrival (DOA) estimator is introduced and the exact Cramer-Rao bound (CRB) is derived for the general case of multiple constant-amplitude polynomial-phase sources. Since the proposed exact ML estimator is computationally intensive, an approximate solution is proposed, originating from the analysis of the ML function in the single chirp signal case. As a result, a new form of spatio-temporal matched filter, referred to as the *chirp beamformer,* is derived, which is applicable to "well separated"' sources that have distinct time-frequency or/and spatial signatures. This beamforming approach requires solving a 3D optimization problem and, therefore, enjoys essentially simpler implementation than that entailed by the exact ML. Simulation results have validated our theoretical CRB analysis and shown the offerings and performance advantages of the proposed estimators. The relationship between the chirp beamformer and well-known conventional narrowband beamforming techniques is delineated.

Publications[2,11,16]

1.11 Blind Source Separation via Joint Anti-Diagonalization

The blind source separation methods can be based on maximum likelihood separation by decorrelation and rotation, neural networks, or on TFDs, which is the focus of this research. The latter has only considered autoterms of the signal TFDs, and exploited the diagonal structure of the source TFDs evaluated at autoterm TF points. We have been able to perform blind source sparation by including, in addition to the autoterms, the crossterms of the source TFDs. This is achieved by exploiting the anti-diagonal structure of the source TFD matrix evaluated at the cross-term TF points. A Jacobi-like algorithm has been derived for the maximization of the Joint-Diagonalization/Joit Anti-Diagonalization (JD/JAD) criterion. The success of the JD or JAD methods depends strongly on the correct selection of the autoterm and crossterm points. Therefore, it is crucial to have a selection procedure that is able to distinguish between autoterm and crossterm points based only on the TFD matrices of the observation. We have proposed an automatic selection procedure to decide, with no a priori knowledge about the sources, whether a considered TF point corresponds to an autoterm or a crossterm. The proposed technique is robust to noise and TF point selection errors and improves the quality of source separation.

Publications[8]

1.12 Spatial evolutionary Spectrum for DOA Estimation and Blind Source Separation

The evolutionary spectrum (ES) was introduced in the sixties by Priestly. This spectrum is based on the modeling of nonstationary signal as a collection of uncorrelated sinusoids with random time-varying amplitudes. The work in this area has lead to the generalization, estimation, and the linkage of ES to TFDs. For processes with restricted time-frequency correlation, referred to as underspread nonstationary random processes, it has already been shown that major definitions of time-varying spectra, such as the generalized Wigner-Ville spectrum and generalized evolutionary spectrum, yield effectively equivalent results. We have successfully combined the concepts of the evolutionary spectrum and array processing. The nonstationary signals received by each sensor of the array will be modeled as a sum of complex sinusoids with time-varying amplitudes. These amplitudes carry information about the direction of arrival. The

time-varying amplitudes using linear estimators based on mean-squared error minimization are first estimated. These estimates are then used to generate the time-varying cross-power distributions between the data across the array. Proceeding similar to spatial joint-variable distributions, the time-varying cross-power estimates computed at high SNR time-frequency points are used for angle estimation. Further, we have shown that the spatial evolutionary spectrum can be directly used for blind source separation. Due to their attractive cross-terms properties, the spatial evolutionary spectrum performance for direction finding and signal recovery compares and potentially exceeds that of TFDs. The same argument applies to positive time-frequency distributions. Next year, we aim to develop positive spatial joint-distributions and examine the offering of positive spectra in angle estimation.

Publications[23]

1.13 Nonstationary Interference Using Subspace Projection Techniques

Combined spatial and time-frequency signatures of signal arrivals at a multi-sensor array are used for nonstationary interference suppression in broadband communication platforms. We have focused on direct-sequence spread-spectrum (DS/SS) communications and shown that with random PN spreading code and deterministic nonstationary interferers, the use of antenna arrays offers increased DS/SS signal dimensionality relative to the interferers. Interference mitigation through *spatio-temporal* subspace projection technique leads to reduced DS/SS signal distortion and improved performance over the case of a single antenna receiver. The angular separation between the interference and desired signals has been shown to play a fundamental role in trading off the contributions of the spatial and t-f signatures to the interference mitigation process.

The main purpose of our work under this topic is to integrate spatial and temporal processing for suppression of nonstationary interferers in DS/SS systems. Specifically, we have successfully extended the projection-based interference mitigation methods to multi-sensor array receivers. Receiving nonstationary interferers at different points in space provides interference suppression techniques with the capability to suppress the interferers using individual or combined time, frequency, and space variables. The proposed multi-sensor interference excision technique builds on the offerings of quadratic TFDs for estimation of 1) the time-frequency subspace and time-frequency signature of nonstationary signals, and 2) the spatial signature of nonstationary sources using direction finding and blind source separations. With the knowledge of the time-frequency and spatial signatures, the objective is to effectively suppress strong nonstationary interferers with few array sensors. The proposed technique does not require the knowledge of the array response or channel estimation of the DS/SS signal, but it utilizes the distinction in both of its spatial- and time-frequency signatures from those of the interferers that impinge on the array. With the combined spatial-time frequency signatures, the projection of the data vector onto the subspace orthogonal to that of the interferers leads to improved receiver performance over that obtained using the subspace projection in the single-sensor case.

Publications[1,22,34]

1.14 Blind Spatial Processing For Frequency Diversity Spread Spectrum Communications: Partial Jamming Suppression

Frequency diversity spread spectrum (FD-SS) has been recently shown to be a powerful tool for digital detection as well as an effective alternative to the traditional spread spectrum techniques, namely direct sequence (DS-SS) and frequency hopping (FH-SS). In a general

context, diversity is conceived by the existence of several replicas (either in code, time space, or frequency). When diversity is available to the receiver either by the structure of the transmitted signal or the architecture of the receiver, optimum spatial signal processing, which is blind to the temporal signal characteristics, can be derived.

We have devised a novel technique to obtain optimum blind spatial processing for frequency diversity spread spectrum (FD-SS) communication systems. The sufficient statistics for a linear combiner, which prove ineffective due to the interferer's spectral location, are modified to yield improved detection under partial jamming in the spectral domain. Robustness to partial time jamming is achieved by extending the notion of replicas over the frequency axis to repetition over the time variable. Analysis and simulations are provided showing the advantages of using FD-SS with spatial diversity to combat interference which is confined to a narrow frequency or time support relative to the desired signal extent in either domain.

Publications[25,30]

1.15 Adaptive Array Processing for Multipath Fading Mitigation Via Exploitation of Filter Banks

The spatial-temporal equalization can be achieved by space-time adaptive processing (STAP) to effectively mitigate inter-symbol interference (ISI) and co-channel interference (CCI). Such a scheme consists of an integrated adaptive array and a temporal equalizer to perform jointly optimum spatial and temporal signal processing. However, solving both the CCI and ISI problems simultaneously by conventional STAP methods is difficult. The recent STAP methods require either large scale matrix inversion, recursive calculation or cascaded CCI and ISI cancellers.

We have proposed an efficient subband adaptive array processing method that utilizes filter banks to mitigate both the CCI and ISI effects in land mobile communications. A subband adaptive array has, in effect, the same function as an STAP system, while the implementation is much easier. The subband adaptive array provides sub-optimal performance in the context of frequency-spatial signal processing, which enhances the signal correlation between multipath rays prior to processing. In subband adaptive arrays, the frequency band of the received signal is divided into smaller bands through the use of filter banks. Analysis filters yield a significant increase in the signal correlation between the multipath rays within each subband. Such an increase is blind in the sense that it does not require a priori knowledge of the arriving signals. As a result of increased correlation, the multipath fading associated with both the desired and the interference signals is reduced, thus yielding faster convergence of the adaptive weight vector as well as fewer degrees-of-freedom (DOF's) required for adequate equalization. The importance to properly choose the analysis and synthesis filters to obtain good equalization performance by subband signal processing is emphasized. The signal correlation enhancements of multipath signals using different filter banks are compared and several simulation examples are provided to confirm the effectiveness of the proposed method.

Publications[3,15,41,44]

2. LIST OF PUBLICATIONS

[1] Y. Zhang and M. Amin, "Array processing for nonstationary interference suppression in DS/SS communications using subspace projection techniques," IEEE Transactions on Signal Processing, vol. 49, no. 12, Dec. 2001.

[2] A. Gershman, M. Pesavento, and M. Amin, "Estimating parameters of multiple wideband polynomial-phase sources in sensor arrays," IEEE Transactions on Signal Processing, vol. 49, no. 12, Dec. 2001.

[3] Y. Zhang, K. Yang, M. Amin, and Y. Karasawa, "Performance analysis of subband arrays," IEICE Transactions on Communications, vol. E-84B, no.9, pp. 2507-2515, Sept. 2001.

[4] M. Amin, W. Mu, and Y. Zhang, "Improved time-frequency synthesis using array manifolds," Proceedings of the IEEE International Symposium on Signal Processing and Its Applications, Kuala-Lampur, Malaysia, Aug. 2001.

[5] Y. Zhang, K. Yang, and M. Amin, "Convergence performance of subband arrays for spatio-temporal equalization," Proceedings of the IEEE Workshop on Statistical Signal Processing, Singapore, pp. 544-547, Aug. 2001.

[6] M. Amin, W. Mu, and Y. Zhang, "Spatial and time-frequency signature estimation of nonstationary sources," Proceedings of the IEEE Workshop on Statistical Signal Processing, Singapore, pp. 313-316, Aug. 2001.

[7] W. Mu, Y. Zhang, and M. Amin, "Bilinear signal synthesis in array processing," Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, May 2001.

[8] A. Belouchrani, K. Abdel Meraim, M. Amin, and A. Zoubir, "Joint anti-diagonalization for blind source separation," Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, Utah, May 2001.

[9] Y. Zhang, W. Mu, and M. Amin, "Subspace analysis of spatial time-frequency distribution matrices," IEEE Transactions on Signal Processing, vol. 49, no. 4, pp. 747-759, April 2001.

[10] Y. Zhang, K. Yang, and M. Amin, "Adaptive array processing for multipath fading mitigation via exploitation of filter banks," IEEE Transactions on Antennas and Propagation, vol. 49, no. 4, pp. 505-516, April 2001.

[11] A. Gershman and M. Amin, "High resolution sensor array processing in the presence of wideband chirps," Proceedings of the 34th Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, Nov. 2000.

[12] M. Amin and Y. Zhang, "Direction finding based on spatial time-frequency distribution matrices," Digital Signal Processing, vol. 10, no. 4, pp. 325-339, Oct. 2000.

[13] Y. Zhang and M. Amin, "Spatial averaging of time-frequency distributions for signal recovery in uniform linear arrays," IEEE Transactions on Signal Processing, vol. 48, no. 10, pp. 2892-2902, Oct. 2000.

[14] W. Mu, M. G. Amin, and Y. Zhang, "Bilinear signal synthesis in array processing," submitted to IEEE Transactions on Signal Processing, Sept. 2000.

[15] Y. Zhang, K. Yang, and M. Amin, "Transform domain array processing for CDMA systems," Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, Pocono Manor, PA, pp. 23-27, Aug. 2000.

[16] A. Gershman, M. Pesavento, and M. G. Amin, "Estimating the parameters of multiple wideband chirp signals in sensor arrays," Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, pp.467-471, Pocono Manor Inn, Pocono Manor, PA, August 2000.

[17] Y. Zhang, W. Mu, and M. Amin, "Time-frequency maximum likelihood methods for direction finding," Journal of the Franklin Institute, vol. 337, no. 4, pp. 483-497, July 2000.

[18] M. Amin, A. Belouchrani, and Y. Zhang, "The spatial ambiguity function and its applications," IEEE Signal Processing Letters, vol. 7, no. 6, pp. 138-140, June 2000.

[19] A. Gershman and M. Amin, "Wideband direction of arrival estimation of multiple chirp signals using spatial time-frequency distributions," IEEE Signal Processing Letters, vol.7, no.6, pp.152-155, June 2000.

[20] A. Gershman and M. Amin, "Coherent wideband DOA estimation of multiple FM signals using spatial time-frequency distributions," Proceedings of the IEEE International Conf. on Acoustics, Speech, and Signal Processing, vol.5, pp.3065-3068, Istanbul, Turkey, June 2000.

[21] Y. Zhang and M. Amin, "Blind separation of sources based on their time-frequency signatures," Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Istanbul, Turkey, pp. 3132-3135, June 2000.

[22] Y. Zhang and M. Amin, "Blind beamforming for suppression of instantaneously narrowband signals in DS/SS communications using subspace projection techniques," Proceedings of the SPIE: Digital Wireless Communication II, vol. 4045, Orlando, FL, pp. 100-109, April 2000.

[23] S. Kayhan and M. Amin, "The application of spatial evolutionary spectrum to direction finding and blind source separation," IEEE Transactions on Signal Processing, April 2000.

[24] M. Amin and Y. Zhang, "Effects of crossterms on the performance of time-frequency MUSIC," Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop, Cambridge, MA, pp. 479-483, March 2000.

[25] M. Lagunas, A. Perez, and M. Amin, "Spatial processing for frequency diversity schemes," IEEE Transactions on Signal Processing, vol.48, no.2, pp.353-362, February 2000.

[26] A. Belouchrani and M. Amin, "A two sensor array blind beamformer for direct sequence spread spectrum communications," IEEE Transactions on Signal Processing, vol.47, no.8, pp.2191-2199, August 1999.

[27] Y. Zhang, W. Mu, and M. Amin, "Maximum likelihood methods for array processing based on time-frequency distributions", Proceedings of the SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations IX, Denver, CO, pp. 502-513, July 1999.

[28] A. Belouchrani and M. Amin, "Time-frequency MUSIC," IEEE Signal Processing Letters, vol.6, no.5, pp.109-110, May 1999.

[29] M. Amin, "Spatial time-frequency distributions for direction finding and blind source separation," Proceedings of SPIE on Aerospace/Defense Sensing, Simulation, and Controls, Orlando, FL, April 1999.

[30] M. Lagunas and Dr. A. Perez, and M. Amin, "Optimum blind spatial processing for frequency diversity spread spectrum communications," Proceedings of SPIE on Aerospace/Defense Sensing, Simulation, and Controls, Orlando, FL, April 1999.

[31] Y. Zhang and M. Amin, "Spatial averaging of time-frequency distributions", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Phoenix, AZ, pp. 1337-1340, March 1999.

[32] M. Amin, "Applications of spatial time-frequency distributions to high resolution radar systems," Proceedings of the NATO RTO-SET Symposium on High Resolution Radar Techniques, Granada, Spain, March 1999,

[33] Bhalla, A. Belouchrani, and M. Amin, "Application of blind signal separation techniques to airborne antenna arrays," Journal of Franklin Institute, January 1999.

[34] M. Amin and G. Mandapati, "Nonstationary interference excision in spread spectrum communications using projection filtering methods," Proceedings of the Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, pp. 827-831, November 1998.

[35] Y. Zhang and M. Amin, "Beamspace time-frequency MUSIC with application to airborne antenna arrays", Proceedings of the 32nd Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, pp. 838-842, Nov. 1998.

[36] M. Amin and W. Williams, "High spectral resolution time-frequency distribution kernels," IEEE Transactions on Signal Processing, vol.46, no.10, pp.2796-2804, October 1998.

[37] A. Belouchrani and M. Amin, "On the use of spatial time frequency distributions for signal extraction," Multidimensional Systems and Sign Processing, Kluwer Academic Publishers, October 1998.

[38] M. Amin and G. Venkatesan, "Trigonometric decomposition of time-frequency distribution kernels," Proceedings of the IEEE Int. Symposium on Time-Frequency and Time-Scale, pp.653-656, Pittsburgh, PA, October 1998.

[39] M. Amin and A. Belouchrani, "Blind source separation using spatial ambiguity functions," Proceedings of the IEEE Int. Symposium on Time-Frequency and Time-Scale, pp.413-416, Pittsburgh, PA, October 1998.

[40] M. Amin, "Time-frequency distributions in statistical signal and array processing," in Highlights of Statistical Signal and Array processing, Celebrating a half Century of Signal Processing, editor A. Hero, IEEE Signal Processing Magazine, vol.15, no.5, September 1998.

[41] Y. Zhang, K. Yang, and M. Amin, "Performance analysis of subband adaptive arrays in multipath propagation environment", Proceedings of the IEEE Signal Processing Workshop on Statistical Signal and Array Processing, Portland, OR, pp. 17-20, Sept. 1998.

[42] Belouchrani and M. Amin, "A two sensor blind beamformer for direct sequence spread spectrum communications," Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, Portland, OR, pp. 120-123, Sept. 1998.

[43] M. Amin and A. Belouchrani, "Time-frequency MUSIC--an array signal processing method based on time-frequency signal representation," Proceedings of the SPIE, Radar Processing, Technology, and Applications III, San Diego, CA, pp. 186-194, July 1998.

[44] Y. Zhang, K. Yang, and M. Amin, "Adaptive subband arrays for multipath fading mitigation", Proceedings of the 1998 IEEE Antennas and Propagation Society International Symposium, Atlanta, GA, pp. 380-383, June 1998.

3. FUNDAMENTAL OFFERINGS OF TIME-FREQUENCY IN ARRAY PROCESSING

Time-frequency distributions have been shown to be a powerful tool in nonstationary signal analysis. So far, most of the work on this subject has focused on temporal signal processing without much attention given to the spatial variable.

To introduce the spatial time-frequency distribution, we first recall that Cohen's class of time-frequency distribution (TFD) of a signal $x(t)$ is given by

$$D_{xx}(t,f)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(t-u,\tau)x(u+\tau/2)x^{*}(u-\tau/2)e^{-j2\pi f\tau}dud\tau \tag{1}$$

where t and f represent the time index and the frequency index, respectively. The kernel $\phi(t,\tau)$ is a function of the time and lag variables. The cross-TFD of two signals $x_1(t)$ and $x_2(t)$ is

$$D_{x_1x_2}(t,f)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(t-u,\tau)x_1(u+\tau/2)x_2^{*}(u-\tau/2)e^{-j2\pi f\tau}dud\tau \tag{2}$$

Expressions (1) and (2) are now used to define the following data spatial time-frequency distribution (STFD). Replacing $x(t)$ in (1) by the data vector $\mathbf{x}(t)$ yields

$$\mathbf{D}_{xx}(t,f)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(t-u,\tau)\mathbf{x}(u+\tau/2)\mathbf{x}^{H}(u-\tau/2)e^{-j2\pi f\tau}dud\tau \tag{3}$$

where $[\mathbf{D}_{xx}(t,f)]_{ij}=D_{x_ix_j}(t,f)$, for $i,j=1,2,....,n$, and the superscript "H" denotes the complex conjugate transpose of a matrix or a vector. The following narrowband model

$$\mathbf{x}(t)=\mathbf{A}\mathbf{s}(t)+\mathbf{n}(t)=\mathbf{y}(t)+\mathbf{n}(t) \tag{4}$$

is assumed, where $\mathbf{y}(t)$ is the signal component of $\mathbf{x}(t)$ and the m x n spatial matrix $\mathbf{A}$ is the mixing matrix. The elements of the m x 1 vector $\mathbf{x}$(t), which represents the measured or sensor data, are multicomponent signals, while the elements of the n x 1 vector $\mathbf{s}(t)$ are often monocomponent signals. $\mathbf{n}(t)$ is an additive noise vector whose elements are temporally and spatially white, zero mean Gaussian distributed process with variance σ^2. Due to the linear data model, the STFD takes the following structure

$$\mathbf{D}_{xx}(t,f)=\mathbf{D}_{yy}(t,f)+\mathbf{D}_{yn}(t,f)+\mathbf{D}_{ny}(t,f)+\mathbf{D}_{nn}(t,f). \tag{5}$$

Under the uncorrelated signal and noise assumption and the zero-mean noise property, $\mathbf{E}[\mathbf{D}_{yn}(t,f)]=E[\mathbf{D}_{ny}(t,f)]=0$, and it follows

$$\overline{\overline{\mathbf{D}}}_{xx}(t,f)=\mathbf{E}[\mathbf{D}_{xx}(t,f)]=\mathbf{D}_{yy}(t,f)+\mathbf{D}_{nn}(t,f)=\mathbf{A}\mathbf{D}_{ss}(t,f)\mathbf{A}^{H}+\sigma^{2}\mathbf{I} \tag{6}$$

where $\overline{\overline{\mathbf{D}}}_{\mathbf{xx}}(t,f)$ is the spatial time-frequency spectrum and $\mathbf{D}_{\mathbf{ss}}(t,f)$ is the TFD of **s**(t). We note that $\overline{\overline{\mathbf{D}}}_{\mathbf{xx}}(t,f)$ is a matrix of dimension m x m, whereas $\mathbf{D}_{\mathbf{ss}}(t,f)$ is of dimension n x n. For narrowband array signal processing applications, **A** holds the spatial information and maps the auto- and cross-TFDs of the source signals into auto- and cross-TFDs of the data.

Expression (6) is similar to that which has been commonly used in blind source separation and direction of arrival (DOA) estimation problems, relating the signal correlation matrix to the data spatial correlation matrix. Here, these correlation matrices are replaced by spatial time-frequency distribution matrices. This means that we can solve these problems in various applications using a new formulation which is more tuned to nonstationary signal environments.

The two subspaces spanned by the principle eigenvectors of $\overline{\overline{\mathbf{D}}}_{\mathbf{xx}}(t,f)$ and the columns of **A** are, therefore, identical. Since the off-diagonal elements are cross-terms of $\mathbf{D}_{\mathbf{ss}}(t,f)$, then this matrix is diagonal for all (t-f) points which correspond only to the signal autoterms. In practice, to simplify the selection of such points of true high power localization, we apply the smoothing kernel $\phi(t,\tau)$ that may significantly decrease the contribution of the cross-terms in the t-f plane.

To our knowledge, there are five key advantages of array processing using time-frequency distributions which have not yet been properly presented and fully utilized. In order to clearly explain these advantages, we use the pictorial diagram in Fig.1. Two sources A and B are incident on a multisensor array. Source A occupies the time-frequency region R_a, whereas source B occupies the time-frequency region R_b. The time-frequency signatures of the two sources overlap, but each source still has a time-frequency region that is not intruded over by the other source. We will assume that the background noise is white.

1) Equation (6) can be easily derived for any arbitrary joint-variables. Time and frequency are indeed the two most commonly used and physically understood parameters. However, by replacing the spatial time-frequency distributions by spatial arbitrary joint-variable distributions, one can relate the sensor joint-variable distributions to the source joint-variable distributions through the same mixing matrix **A**. As shown below, there are situations where it is preferable to consider other domains such as the ambiguity domain, where the locations of the signals and their cross-terms are guided by properties and mechanisms different than those associated with the time-frequency domain.

2) Equation (6) is valid for all time-frequency points. The inability to perform ensemble averaging gives rise to the question of whether one time-frequency point suffices for adequate direction finding and source separation, and how sensitive the performance is to a random choice of a t-f point? Further, if several t-f points are used, then how to choose and combine these points for improved performance, and whether the method of combining should differ depending on the task in hand? Direction finding techniques require $\mathbf{D}_{\mathbf{ss}}(t,f)$ to be full rank, preferably diagonal. Some blind source separation techniques demand the diagonal structure of the same matrix without degenerate eigenvalues. These properties along with high SNR requirements may be difficult to achieve using a single time-frequency point. We have identified two different methods to integrate several t-f points into equation (6). One method is based on a simple

averaging performed over parts or the entire time-frequency regions of the signals of interest. The second method incorporates desirable time-frequency points into joint diagonalization or joint block diagonalization schemes. Both methods aim to fully utilize the points of maximum power concentration and avoid the time-frequency region of significant noise contamination.

3) The time-frequency distribution of the white noise is distributed all over the time-frequency domain, whereas the TFDs of the source and jammer waveforms are likely to be confined to much smaller regions. Referring to Fig.1, the noise is spread over both R_a and R_b as well as the complement region R_c. If the time-frequency points (t, f) used in either the averaging or joint diagonalization approaches belong the noise only region R_c, then no information of the incident waveforms is used and, as such, no reasonable source localization and signal separation outcomes can be obtained. Accordingly, the performance in this case is expected to be worse than conventional approaches. On the other hand, if all points (t, f) in Fig.1 are used, and the employed TFD satisfies the marginal constraints such as the Wigner distribution, then it is easily shown that only the average power is considered. As a result, the problem simplifies to the second order covariance based matrix approach, traditionally used in high resolution direction of arrival estimation. This is an important property, as it casts the conventional techniques as special cases of the proposed framework based on time-frequency analysis. Finally, if we confine the (t, f) points to R_a and R_b, then only the noise part in these regions is included. The result of leaving out the points (t, f) which are not part of the time-frequency signatures of the signal arrivals is, in essence, enhancing the SNR of the input to be utilized by the source localization and signal separation techniques.

4) By only selecting (t, f) points which belong to the t-f signature of one source, then this source will be the only one considered by equation (6). This is, in essence, equivalent to implicitly performing spatial filtering to remove other sources from consideration. It is important to note that such removal does not come at the expense of reduction of the number of degrees of freedom, as it is the case in beamspace processing, but the problem remains a sensor space processing with the original number of degrees of freedom is kept intact. This represents a key contribution of TFDs to the direction finding and angle estimation area. An antenna array can, indeed, be used to localize a number of sources equal or even greater than its number of sensors. The fundamental condition is that there must be time-frequency regions over which the respective time-frequency signatures of the sources do not overlap. In principle, the lower limit on the size of such regions is a single time-frequency point. Referring to Fig.1 and considering the case of two sensors, if all t-f points incorporated in direction finding belong to region R_a and not R_b, then the signal subspace defined by equation (6) is one-dimensional. In effect, by excluding source B, a one-dimensional noise subspace is established, which allows us to proceed with noise-subspace based high resolution techniques for localization of source A. Within the proposed framework, one can localize one source at a time or a set of selected sources, depending on the array size, overlapping and distinct time-frequency regions, and the dimension of the noise subspace necessary to achieve the required resolution performance. The same concepts and advantages of t-f point selection discussed above for direction finding can be applied to blind source separations.

5) The a priori knowledge of some temporal characteristics or the nature of time-varying frequency contents of the sources may permit direct selection of the t-f regions used in equation (6). In general, if we choose a joint-variable domain, where a class of signals is confined to a specific known joint-variable region, then one can perform direction finding and source separation for only this specific class. For instance, it is known that in the ambiguity domain all

fixed frequency sinusoids map to the vertical axis, independent of their amplitudes, frequencies, and phases are. By only incorporating the points on the vertical axis, which represents the time-lag variable, we focus on separating and localizing narrowband components in the presence of broadband signals.

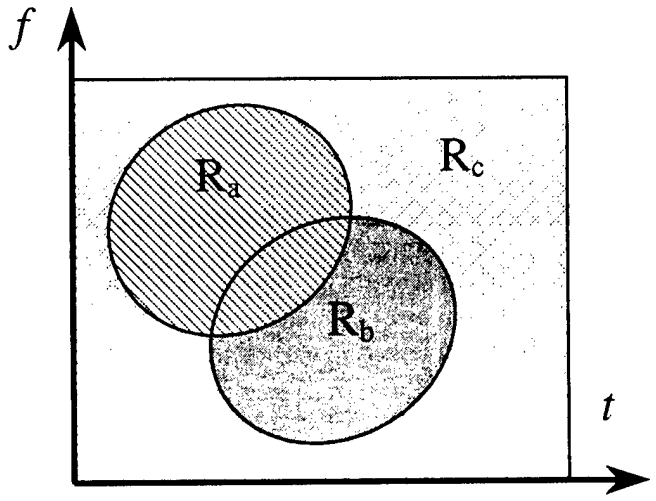


Fig. 1 Signals with different time-frequency signature

4. PUBLICATIONS

- Array processing for nonstationary interference suppression in DS/SS communications using subspace projection techniques
- Estimating parameters of multiple wideband polynomial-phase sources in sensor arrays
- Performance analysis of subband arrays
- Joint anti-diagonalization for blind source separation
- Subspace analysis of spatial time-frequency distribution matrices
- Spatial averaging of time-frequency distributions for signal recovery in uniform linear arrays
- Bilinear signal synthesis in array processing
- Time-frequency maximum likelihood methods for direction finding
- The spatial ambiguity function and its applications

Array processing for nonstationary interference suppression in DS/SS communications using subspace projection techniques

Yimin Zhang and Moeness G. Amin
Department of Electrical and Computer Engineering,
Villanova University, Villanova, PA 19085
E-mail: yimin@ieee.org, moeness@ece.villanova.edu

Abstract

Combined spatial and time-frequency signatures of signal arrivals at a multi-sensor array are used for nonstationary interference suppression in direct-sequence spread-spectrum (DS/SS) communications. With random PN spreading code and deterministic nonstationary interferers, the use of antenna arrays offers increased DS/SS signal dimensionality relative to the interferers. Interference mitigation through spatio-temporal subspace projection technique leads to reduced DS/SS signal distortion and improved performance over the case of a single antenna receiver. The angular separation between the interference and desired signals is shown to play a fundamental role in trading off the contribution of the spatial and time-frequency signatures to the interference mitigation process. The expressions of the receiver SINR implementing subspace projections are derived and numerical results are provided.

Keywords

Time-frequency distribution, interference suppression, subspace projection, array processing, direct-sequence spread-spectrum communication.

This work is supported by the Office of Naval Research under Grant No. N00014-98-1-0176, and the Air Force Research Laboratory under Grant No. F30602-00-1-0515.

I. Introduction

There are several methods that have been proposed for interference suppression in direct-sequence spread-spectrum (DS/SS) communications, most have been related to one domain of operation [1], [2]. These methods include the narrowband interference waveform estimation [3], [4], frequency domain interference excision [5], zero-forcing techniques [6], adaptive subspace-based techniques [7], [8], and minimum-mean-square error (MMSE) interference mitigation techniques [9].

Nonstationary interferers, which have model parameters that change with time, are particularly troublesome due to the inability of a single domain mitigation algorithm to adequately remove their effects. The recent development of the quadratic time-frequency distributions (TFDs) for improved signal power localization in the time-frequency plane has motivated several new approaches for excision of interference with rapidly time-varying frequency characteristics in the DS/SS communication systems. Comprehensive summary of TFD-based interference excision is given in reference [10]. The two basic methods for time-frequency excision are based on notch filtering and subspace projections. Utilization of the interference instantaneous frequency (IF), as obtained via TFDs, to design an open loop adaptive notch filter in the temporal domain, has been thoroughly discussed in [11], [12]. Subspace projection methods, commonly used for mitigating narrowband interference [13], [14], have been recently introduced for suppression of frequency modulated (FM) interference and shown to properly handle multi-component interference, reduce the self-noise, and improve the receiver performance beyond that offered by other time-frequency based techniques [15], [16], [17].

The main purpose of this paper is to integrate spatial and temporal processing for suppression of nonstationary interferers in DS/SS communication systems. Specifically, we extend the projection-based interference mitigation techniques in [15], [16], [17] to multi-sensor array receivers. The proposed multi-sensor interference excision technique

builds on the offerings of quadratic time-frequency distributions for estimation of 1) the time-frequency subspace and time-frequency signature of nonstationary signals, and 2) the spatial signature of nonstationary sources using direction finding and blind source separations. With the knowledge of the time-frequency and spatial signatures, the objective is to effectively suppress strong nonstationary interferers with few array sensors. The proposed technique does not require the knowledge of the array response or channel estimation of the DS/SS signal, but it utilizes the distinction in both of its spatial- and time-frequency signatures from those of the interferers that impinge on the array. With the combined spatial-time-frequency signatures, the projection of the data vector onto the subspace orthogonal to that of the interferers leads to improved receiver performance over that obtained using the subspace projection in the single-sensor case.

The rest of the paper is organized as follows. In Section II, the signal model is described. Section III briefly reviews the subspace projection technique. We present in Section IV blind beamforming based on subspace projection and derive the receiver output signal-to-interference-plus-noise ratio (SINR). Several numerical results are given in Section V. Section VI concludes this paper.

II. Signal Model

In DS/SS communications, each symbol is spread into $L = T/T_c$ chips, where T and T_c are, respectively, the symbol duration and chip duration. We use discrete-time form, where all signal arrivals are sampled at the chip-rate of the DS/SS signal. The symbol-rate source signal is denoted as $s(n)$, and the aperiodic binary spreading sequence of the nth symbol period is represented by $c(n,l) \in \pm 1, l = 0, 1, \cdots, L-1$. The chip-rate sequence of the DS/SS signal can be expressed as

$$d(k) = s(n)c(n,l) \qquad \text{with} \quad k = nL + l. \tag{1}$$

For notation simplicity, we use $c(l)$ instead of $c(n,l)$ for the spreading sequence.

We consider an antenna array of N sensors. The propagation delay between antenna elements is assumed to be small relative to the inverse of the transmission bandwidth, so that the received signal at the N sensors are identical to within complex constants. The received signal vector of the DS/SS signal at the array is expressed by the product of the chip-rate sequence $d(k)$ and its spatial signature $\underline{\mathrm{h}}$,

$$\underline{\mathrm{x}}_s(k) = d(k)\underline{\mathrm{h}}. \tag{2}$$

The channel is restricted to flat-fading, and is assumed fixed over the symbol length, and as such $\underline{\mathrm{h}}$ in the above equation is not a function of k.

The array vector associated with a total of U interference signals is given by

$$\underline{\mathrm{x}}_u(k) = \sum_{i=1}^{U} \underline{\mathrm{a}}_i u_i(k) \tag{3}$$

where $\underline{\mathrm{a}}_i$ is the array response to the ith interferer, $u_i(k)$. Without loss of generality, we set $||\underline{\mathrm{h}}||_F^2 = N$ and $||\underline{\mathrm{a}}_i||_F^2 = N$, $i = 1, 2, \cdots, U$, where $||\cdot||_F$ is the Frobenius norm of a vector. The input data vector is the sum of three components,

$$\underline{\mathrm{x}}(k) = \underline{\mathrm{x}}_s(k) + \underline{\mathrm{x}}_u(k) + \underline{\mathrm{b}}(k) = d(k)\underline{\mathrm{h}} + \sum_{i=1}^{U} \underline{\mathrm{a}}_i u_i(k) + \underline{\mathrm{b}}(k) \tag{4}$$

where $\underline{\mathrm{b}}(k)$ is the additive noise vector. In regards to the above equation, we make the following assumptions.

A1) The information symbols $s(n)$ is a wide-sense stationary process with $E[s(n)s^*(n)] = 1$, where $E[\cdot]$ is the statistical expectation operator, and the superscript * denotes complex conjugation. The spreading sequence $c(k)$ is a binary random sequence with $E[c(k)c(k+l)] = \delta(l)$, where $\delta(l)$ is the delta function.[1]

A2) The noise vector $\underline{\mathrm{b}}(k)$ is zero-mean, temporally and spatially white with

$$E[\underline{\mathrm{b}}(k)\underline{\mathrm{b}}^T(k+l)] = \mathbf{0} \quad \text{for all } l,$$

[1]This assumption is most suitable for military applications and P-code GPS.

and

$$E[\underline{b}(k)\underline{b}^H(k+l)] = \sigma\delta(l)\mathbf{I}_N,$$

where σ is the noise power, the superscripts ^{T} and ^{H} denote transpose and conjugate transpose, respectively, and $\mathbf{I}_N$ is the $N \times N$ identity matrix.

A3) The signal and noise are statistically uncorrelated.

III. Subspace Projection

The aim of subspace projection techniques is to remove the interference components before despreading by projecting the input data on the subspace orthogonal to the interference subspace, as illustrated in Fig. 1. The receiver block diagram is shown in Fig. 2.

A nonstationary interference, such as an FM signal, often shares the same bandwidth with the DS/SS signal and noise. As such, for a chirp signal or a signal with high-order frequency laws, the signal spectrum may span the entire frequency band, and the sample data matrix loses its complex exponential structure responsible for its singularity. Therefore, the interference subspace can no longer be obtained from the eigendecomposition of the sample data matrix [13], [15] or the data covariance matrix [14], as it is typically the case in stationary environments. The nonstationary interference subspace, however, may be constructed using the interference time-frequency signature. Methods for estimating the instantaneous frequency, instantaneous bandwidth, and more generally, a time-frequency subspace, based on the signal time-frequency localization properties are, respectively, discussed in references [15], [18], [19].

For the general class of FM signals, and providing that interference suppression is performed separately over the different data symbols, the interference subspace is one-dimensional in an L-dimensional space. We note that since an FM interference has a constant amplitude, its respective data vector can be determined from the IF up to a complex multiplication factor. The unit norm normalization of this vector represents the one-dimensional interference subspace basis vector. Among candidate methods of IF es-

timation is the one based on the time-frequency distributions. For example, the discrete form of Cohen's class of TFD of a signal $x(t)$ is given by [20]

$$D_{xx}(t,f) = \sum_{m=-\infty}^{\infty} \sum_{\tau=--\infty}^{\infty} \phi(m,\tau) x(t+m+\tau) x^*(t+m-\tau) e^{-j4\pi f\tau}, \tag{5}$$

where $\phi(m,\tau)$ is a time-frequency kernel that could be signal-dependent. The TFD concentrates the interference signal power around the IF and makes it visible in the noise and pseudo-random (PN) sequence background [18], [21]. It has been shown that, for linear FM signals, Radon-Wigner transform provides improved IF estimates over the TFD [22]. Parametric methods using autoregressive model have also been proposed [23].

Other nonstationary interference with instantaneous bandwidth or spread in the time-frequency domain are captured in a higher-dimension subspace. In this case, the interference subspace can be constructed from the interference localization region Ω in the time-frequency domain (see, for example, [15]). The subspace of interest becomes that which fills out the interference time-frequency region Ω energetically, but has little or no energy outside Ω.

Interference-free DS/SS signals are obtained by projecting the received data vector (in the temporal domain processing, the vector consists of data samples at different snapshots) on the subspace orthogonal to the interference subspace.

A. Temporal Processing

In the single-sensor receiver, the input data is expressed as

$$x(k) = x_s(k) + x_u(k) + b(k) = d(k) + \sum_{i=1}^{U} u_i(k) + b(k). \tag{6}$$

Using L sequential chip-rate samples of one symbol of the received signals at time index

k, we obtain the following input vector

$$\begin{aligned}
&[x(k) \;\; x(k-1) \;\; \cdots \;\; x(k-L+1)]^T \\
&= [x_s(k) \;\; x_s(k-1) \;\; \cdots \;\; x_s(k-L+1)]^T \\
&+ [x_u(k) \;\; x_u(k-1) \;\; \cdots \;\; x_u(k-L+1)]^T \\
&+ [b(k) \;\; b(k-1) \;\; \cdots \;\; b(k-L+1)]^T
\end{aligned} \tag{7}$$

or simply

$$X(k) = X_s(k) + X_u(k) + B(k). \tag{8}$$

We drop the variable k for simplicity, with the understanding that processing is performed over the nth symbol that starts at the kth chip. Then, equation (8) becomes

$$X = X_s + X_u + B. \tag{9}$$

Below, we relax the FM condition used in [13], [16] that translates to a single dimension interference. The general case of an interference occupying higher dimension subspace is considered. We assume that the ith interferer spans M_i dimensional subspace, defined by the orthonormal basis vectors, $V_{i,1}, V_{i,2}, \cdots, V_{i,M_i}$, and the different interference subspaces are disjoint. Define

$$V_i = [V_{i,1} \; V_{i,2} \; \cdots \; V_{i,M_i}] \tag{10}$$

and let $M = \sum_{i=1}^{U} M_i$ as the number of total dimensions of the interferers. With $L > M$, the $L \times M$ matrix

$$V = [V_1 \; V_2 \; \cdots \; V_U], \quad V_i \bigcap V_j = \Phi \quad \text{for } i \neq j \tag{11}$$

is full rank and its columns span the combined interference subspace J. The respective projection matrix is

$$\overline{P} = V \left(V^H V\right)^{-1} V^H. \tag{12}$$

The projection matrix associated with the interference orthogonal subspace, G, is then given by

$$P = \mathbf{I}_L - V(V^H V)^{-1} V^H. \tag{13}$$

When applied to X, matrix P projects the input data vector onto G, and results in

$$X_\perp = PX = PX_s + PB, \tag{14}$$

which no longer includes any interference component.

The single-sensor receiver implementing subspace projection for excision of a single instantaneously narrowband FM interferer (i.e., $U = 1, M_1 = 1$) in DS/SS communications is derived in [24]. The receiver SINR is shown to be

$$\text{SINR} = \frac{(L-1)^2}{\left(1 - \frac{2}{L}\right) + \sigma(L-1)} = \frac{L-1}{\frac{L-2}{L(L-1)} + \sigma}. \tag{15}$$

For typical values of L, $(L-2)/(L-1) \approx 1$, and equation (15) can be simplified as

$$\text{SINR} \approx \frac{L-1}{\sigma + 1/L}. \tag{16}$$

Compared to the interference-free environment, where the receiver SINR is L/σ, nonstationary interference suppression in (16) is achieved by reducing the processing gain by 1 and increasing the noise power by the self-noise factor of $1/L$.

IV. Subspace Projection in Multi-Sensor Receiver

In this section, we consider nonstationary interference excision in multi-sensor receivers using subspace projections. We note that if the subspace projection method discussed in Section III is extended to an N-element array by suppressing the interference independently in each sensor data and then combining the results by maximum ratio combining (see Fig. 3), then it is straightforward to show that the receiver SINR is given by

$$\text{SINR} \approx \frac{N(L-1)}{\sigma + N/L}. \tag{17}$$

The above extension, although clearly improves over (16), does not utilize the potential difference in the spatial signatures of signal arrivals, and, therefore, is inferior to the receiver proposed in this Section.

A. Spatio-Temporal Signal Subspace Estimation

To construct the spatio-temporal signal subspace of the interference signals, it is important to estimate both the time-frequency signature (or subspace) and the spatial signature of each interferer. The IF estimation of an FM interference signal based on time-frequency distribution is addressed in Section III. It is noteworthy that when multiple antennas are available, the TFD may be computed at each sensor data separately and then averaged over the array. This method has been shown in [25] to improve the IF estimation, as it reduces noise and crossterms that often obscure the source true power localization in the time-frequency domain.

On the other hand, the estimation of source spatial signature can be achieved, for example, by using direction finding and source separation techniques. When the interference signals have clear bearings, methods like MUSIC [26] and maximum likelihood (ML) [27] can be used to estimate the steering matrix of the interference signals. These methods can be revised to incorporate the TFD of the signal arrivals for improved performance [28], [29], [30]. On the other hand, in fading channels where the steering vector loses its known structure due to multipath, blind source separation methods should be used [31], [32], [33]. Since the interferers in DS/SS communications often have relatively high power, good spatial signature estimation is expected.

More conveniently, the spatial signatures can be simply estimated by using matched filtering once the time-frequency signatures are provided. The maximum likelihood estimator for the vector $\underline{a}_i$ is obtained as

$$\hat{\underline{a}}_i = \sqrt{N} \sum_{k=0}^{L-1} \hat{u}_i^*(k)\underline{x}(k) / || \sum_{k=0}^{L-1} \hat{u}_i^*(k)\underline{x}(k)||_F, \tag{18}$$

where $\hat{u}_i(k)$ is the estimated waveform of the ith interferer. It is noted that the possible phase ambiguity in the waveform estimation of $\hat{u}_i(k)$ does not affect the estimation of the spatial signature. For slowly varying channels, the above average can also be performed over multiple symbols to improve the estimation accuracy.

In the analysis presented herein, we assume knowledge of the interference subspace and its spatial signature to derive the receiver SINR.

B. Proposed Technique

The subspace projection problem for nonstationary interference suppression in DS/SS communications is now considered within the context of multi-sensor array using N array elements. We use one symbol DS/SS signal duration (i.e., L chip-rate temporal snapshots), and stack L discrete observations to construct an $NL \times 1$ vector of the received signal sequence in the joint spatio-temporal domain. In this case, the received signal vector in (4) becomes

$$\begin{aligned}
&\left[\underline{\mathrm{x}}^T(k) \;\; \underline{\mathrm{x}}^T(k-1) \;\; \cdots \;\; \underline{\mathrm{x}}^T(k-L+1)\right]^T \\
&= \left[\underline{\mathrm{x}}_s^T(k) \;\; \underline{\mathrm{x}}_s^T(k-1) \;\; \cdots \;\; \underline{\mathrm{x}}_s^T(k-L+1)\right]^T \\
&+ \left[\underline{\mathrm{x}}_u^T(k) \;\; \underline{\mathrm{x}}_u^T(k-1) \;\; \cdots \;\; \underline{\mathrm{x}}_u^T(k-L+1)\right]^T \\
&+ \left[\underline{\mathrm{b}}^T(k) \;\; \underline{\mathrm{b}}^T(k-1) \;\; \cdots \;\; \underline{\mathrm{b}}^T(k-L+1)\right]^T
\end{aligned} \tag{19}$$

or simply

$$\mathbf{X} = \mathbf{X}_s + \mathbf{X}_u + \mathbf{B}, \tag{20}$$

where again the variable k is dropped for simplicity.

In (19), the interference vector in the single-sensor problem, given by (7), is extended to a higher dimension. With the inclusion of both temporal and spatial samples, the mth basis of the ith interference becomes

$$\mathbf{V}_{i,m} = V_{i,m} \otimes \underline{\mathrm{a}}_i \tag{21}$$

and

$$\mathbf{V}_i = \left[\mathbf{V}_{i,1} \; \mathbf{V}_{i,2} \; \cdots \; \mathbf{V}_{i,M_i}\right], \tag{22}$$

where $\otimes$ denotes the Kronecker product. The columns of the $LN \times M$ matrix

$$\mathbf{V} = \left[\mathbf{V}_1 \; \mathbf{V}_2 \; \cdots \; \mathbf{V}_U\right] \tag{23}$$

spans the overall interference signal subspace.

For independent spatial signatures, the matrix rank is M. The orthogonal projection matrix is given by

$$\mathbf{P} = \mathbf{I}_{LN} - \mathbf{V}\left(\mathbf{V}^H\mathbf{V}\right)^{-1}\mathbf{V}^H. \tag{24}$$

The projection of the signal vector on the orthogonal subspace of the interferers' yields

$$\mathbf{X}_\perp = \mathbf{PX} = \mathbf{PX}_s + \mathbf{PB}. \tag{25}$$

The block diagram of the proposed method is presented in Fig. 4. As shown in the next section, effective interference suppression can be achieved solely based on the spatial signatures or the time-frequency signatures, or it may require both information.

C. Performance Analysis

Below we consider the performance of the multi-sensor receiver system implementing subspace projections. Recall that

$$V_{i,m}^H V_{j,n} = 0 \quad \text{for any} \quad i, m \neq j, n. \tag{26}$$

and

$$\mathbf{V}^H\mathbf{V} = N\mathbf{I}_M, \tag{27}$$

the projection matrix $\mathbf{P}$ becomes

$$\mathbf{P} = \mathbf{I}_{LN} - \frac{1}{N}\mathbf{V}\mathbf{V}^H. \tag{28}$$

The signal vector $\mathbf{X}_s$ can be rewritten as

$$\begin{aligned}
\mathbf{X}_s &= \left[\underline{\mathrm{x}}_s^T(k) \;\; \underline{\mathrm{x}}_s^T(k-1) \;\; \cdots \;\; \underline{\mathrm{x}}_s^T(k-L+1)\right]^T \\
&= \left[d(k)\underline{\mathrm{h}}^T \;\; d(k-1)\underline{\mathrm{h}}^T \;\; \cdots \;\; d(k-L+1)\underline{\mathrm{h}}^T\right]^T \\
&= s(n)\left[c(L-1)\; c(L-2) \;\; \cdots \;\; c(0)\right]^T \otimes \underline{\mathrm{h}} \\
&\underline{\underline{\Delta}}\; s(n)\mathbf{q},
\end{aligned} \tag{29}$$

where the $LN \times 1$ vector

$$\mathbf{q} = \left[c(L-1)\; c(L-2) \;\; \cdots \;\; c(0)\right]^T \otimes \underline{\mathrm{h}} \;\underline{\underline{\Delta}}\; \underline{\mathrm{c}} \otimes \underline{\mathrm{h}} \tag{30}$$

defines the spatio-temporal signature of the desired DS/SS signal. $\mathbf{q}$ is the extension of the DS/SS code by replicating it with weights defined by the signal spatial signature.

By performing despreading and beamforming, the symbol-rate decision variable is given by

$$y(n) = \mathbf{q}^H \mathbf{X}_\perp(k) = s(n)\mathbf{q}^H \mathbf{P}\mathbf{q} + \mathbf{q}^H \mathbf{P}\mathbf{B}(k) \triangleq y_1(n) + y_2(n), \tag{31}$$

where $y_1(n)$ is the contribution of the desired DS/SS signal to the decision variable, and $y_2(n)$ is the respective contribution from the noise.

The SINR of the array output becomes (see Appendix A)

$$\text{SINR} = \frac{E^2[y(k)]}{var[y(k)]} = \frac{\left(L - \sum_{i=1}^{U} M_i|\beta_i|^2\right)^2}{\left(\left(\sum_{i=1}^{U} M_i|\beta_i|^2\right)^2 - 2\sum_{i=1}^{U}\xi_i|\beta_i|^4\right) + \frac{\sigma}{N}\left(L - \sum_{i=1}^{U} M_i|\beta_i|^2\right)}, \tag{32}$$

where ξ_i is defined in (A.9), and β_i is the spatial correlation coefficient between the spatial signatures $\underline{\mathrm{h}}$ and $\underline{\mathrm{a}}_i, i = 1, 2, \cdots, U$, and is given by

$$\beta_i = \frac{1}{N}\underline{\mathrm{h}}^H \underline{\mathrm{a}}_i. \tag{33}$$

Note that when the noise power is small, i.e., $\sigma \ll 1$, the variance of y_1 becomes dominant, and the output SINR reaches the following upper bound

$$\text{SINR}_{\text{high SNR}} \approx \frac{\left(L - \sum_{i=1}^{U} M_i|\beta_i|^2\right)^2}{\left(\sum_{i=1}^{U} M_i|\beta_i|^2\right)^2 - 2\sum_{i=1}^{U}\xi_i|\beta_i|^4}. \tag{34}$$

This result is affected by the factors L, M_i, $|\beta_i|$, and $\xi_i, i = 1, \cdots, U$. On the other hand, when the noise level is very high, i.e., $\sigma \gg 1$, the noise variance plays a key role in determining $var[y(k)]$, and the output SINR becomes

$$\text{SINR}_{\text{low SNR}} \approx \frac{\left(L - \sum_{i=1}^{U} M_i|\beta_i|^2\right)^2}{\frac{\sigma}{N}\left(L - \sum_{i=1}^{U} M_i|\beta_i|^2\right)} = \frac{N}{\sigma}\left(L - \sum_{i=1}^{U} M_i|\beta_i|^2\right). \tag{35}$$

Unlike the high input SNR case, the output SINR in (35) also depends on both N and σ. Comparing (34) and (35), it is clear that the improvement in the receiver SINR becomes more significant when the spatial signatures produce small spatial correlation coefficients and under high SNR.

Next, we consider some specific important cases. When $\beta_i = 0, i = 1, \cdots, U$, $var[y_1(n)] = 0$, the receiver SINR in (32) becomes SINR $= LN/\sigma$. This is to say, the output SINR is improved by a factor of LN over the input signal-to-noise ratio (SNR) (not the input SINR!). This implies that the interferers are suppressed by spatial selectivity of the array and their suppression does not cause any distortion of the temporal characteristics of the DS/SS signal. The DS/SS signal in this case enjoys the array gain that contributes the factor N to the SINR.

For a single FM interferer ($U = 1$, $M_1 = 1$), equation (32) becomes

$$\text{SINR} = \frac{\left(L - |\beta_1|^2\right)^2}{\left(1 - \frac{2}{L}\right)|\beta_1|^4 + \frac{\sigma}{N}\left(L - |\beta_1|^2\right)}. \tag{36}$$

It is easy to show that SINR in (36) monotonously decreases as $|\beta_1|$ increases, and the lower bound of the SINR is reached for $|\beta_1| = 1$, which is the case of the desired DS/SS signal and the interference signal arriving from the same direction. With a unit value of $|\beta_1|$,

$$\text{SINR} = \frac{(L-1)^2}{\left(1 - \frac{2}{L}\right) + \frac{\sigma}{N}(L-1)} \approx \frac{N(L-1)}{\sigma + N/L}. \tag{37}$$

This result is the same as that of the single-sensor case developed in [16], except for the appearance of the array gain, N, for the desired DS/SS signal over the noise. This equation also coincides with (17). That is, the independent multi-sensor subspace projection, illustrated in Fig. 3, results in the same output SINR with the proposed multi-sensor subspace projection method when $|\beta_1| = 1$.

On the other hand, the maximum value in (36) corresponds to $\beta = 0$, and is equal to SINR $= LN/\sigma$, as discussed above. For the illustration of the SINR behavior, we plot in

Fig. 5 the SINR in (36) versus $|\beta_1|$ for a two-sensor array, where $L = 64$, and one FM jammer is considered with $M = 7$. The input SNR is 0dB.

Given the temporal and spatial signatures, the proposed technique simplifies to two consecutive tasks. The first is to estimate the spatio-temporal signature. When using multiple antenna receivers, a basis vector of the orthogonal projection matrix is obtained by the Kronecker product of a jammer's temporal signature and its spatial signature, that results in the $LN \times LN$ orthogonal project matrix instead of $L \times L$ in the single antenna case. The second task is jammer suppression via subspace projection. This involves the multiplication of an $LN \times LN$ matrix and an $LN \times 1$ vector.

Note such increase in computations is natural due to increase of dimensionality. It is noteworthy that array processing expands overall space dimensionality but maintains the jammer subspace dimension. As a result, it yields improved SINR performance over temporal processing or spatial processing only methods.

V. Numerical Results

A two-element array is considered with half-wavelength spacing. The DS/SS signal uses random spreading sequence with $L = 64$. The AOA of the DS/SS signal is 0 degree from broadside ($\theta_D = 0°$).

We consider two interference signals. Each interference signal is assumed to be made up of uncorrelated FM component with $M_i = 7, i = 1, 2$. The overall interference subspace is M=14. The AOAs of the two interferers are $\theta_J = [40°, 60°]$. The respective spatial correlations in this example are $|\beta_1| = 0.53$ and $|\beta_2| = 0.21$. Note that, in the subspace projection method, the output SINR is independent of the input jammer-to-signal ratio (JSR), since the interferers are entirely suppressed, regardless of their input power. Fig. 6 shows the receiver SINR versus the input SNR. The upper bounds correspond to interference-free data. For high input SNR, the receiver SINR is decided by the induced signal distortion, described by the variance given in (A.10). It is evident from Fig. 6 that the two-antenna

receiver outperforms the single-antenna receiver case by a factor much larger than the array gain. Since the output SINR in the two-antenna receiver highly depends on the spatial correlation coefficients, the curves corresponding to a two-sensor array in Fig. 6 will assume different values upon changing β_1, or/and β_2. The best performance is achieved at $\beta_1 = \beta_2 = 0$.

Fig. 7 shows the receiver SINR versus the number of chips per symbol (L). We let L vary from 8 to 4096, whereas the input SNR is fixed at 0 dB. The two interference signals are incident on the array with angles $\theta_J = [40°, 60°]$. They are assumed to maintain their time-frequency spread with increased value of L. As such, the respective dimensions of their subspaces grow proportional to the number of chips per symbol. In this example, the dimension of each interference signal is assumed to be 10 percent of L (round to the nearest integer). The output SINR improvement by performing array processing at different L is evident from this figure. It is seen that, unlike the case of the instantaneously narrowband FM interference, where the output SINR increases rapidly as L increases, the output SINR in the underlying scenario ceases to increase as L assumes large values. This is because the rank of the interference signal subspace increases with L.

In Fig. 8 we investigate the receiver SINR performance versus the number of array sensors. In this figure, L is set at 64, and the input SNR is 0 dB. Two interference signals composed of uncorrelated FM components are considered, and $M_i = 7, i = 1, 2$, are assumed. Two examples are used to examine the effect of different AOAs. In the first example, $\theta_J = [40°, 60°]$. The output SINR improves sharply as the number of array sensors increases from one to three, beyond which the improvement becomes insignificant. The differences in the above AOAs of the desired DS/SS signal and the interference signals are relatively large, and a small number of array sensors leads to negligible spatial correlation coefficients. We also show a case with closely spaced interference signals where $\theta_J = [5°, 15°]$. In this case, the output SINR slowly improves as the number of array sensors increases.

It is noted that, when we consider a specific case, the output SINR does not increase monotonously with the number of array sensors. This is because the relationship between the spatial correlation coefficient and the AOAs is by itself not monotonous. Nevertheless, when we consider the general case with different AOA combinations, high number of array sensors often reduce the spatial correlation coefficients.

VI. Conclusions

In this paper, subspace projection techniques were employed to suppress nonstationary interferers in direct-sequence spread-spectrum (DS/SS) communication systems. Interference suppression is based on the knowledge of both the interference time-frequency and spatial signatures. While the former is based on instantaneous frequency information that can be gained using several methods, including time-frequency distributions, the later can be provided from applying higher resolution methods or blind source separation techniques to the signal arrivals.

The differences between the DS/SS signal and interference signatures both in the time-frequency and spatial domains equip the projection techniques with the ability to remove the interference with a minimum distortion of the desired signal.

The receiver performance based on subspace projections was analyzed. It was shown that the lower performance bound is obtained when the sources have the same angular position. In this case, the problem becomes equivalent to a single-antenna receiver with only the presence of the array gain. On the other hand, the upper bound on performance is reached in the interference-free environment and also corresponds to the case in which the spatial signatures of the interference signals are orthogonal to that of the DS/SS signal.

Numerical results were presented to illustrate the receiver SINR dependency on spatial correlation coefficient, input SNR, and the PN sequence length.

Appendix A

To derive the output SINR expression, we use $s(n) = +1$ (the output SINR is independent of $s(n)$ and same result follows when $s(n) = -1$). Then,

$$\begin{aligned} E[y_1(n)] &= E\left[\mathbf{q}^H \mathbf{P} \mathbf{q}\right] \\ &= E\left[\mathbf{q}^H (\mathbf{I} - \frac{1}{N}\mathbf{V}\mathbf{V}^H)\mathbf{q}\right] \\ &= E\left[\mathbf{q}^H \mathbf{q}\right] - \frac{1}{N} E\left[\mathbf{q}^H \mathbf{V}\mathbf{V}^H \mathbf{q}\right] \\ &= LN - \frac{1}{N} E\left[\mathbf{q}^H \sum_{i_1=1}^{U} \mathbf{V}_{i_1} \sum_{i_2=1}^{U} \mathbf{V}_{i_2}^H \mathbf{q}\right] \\ &= LN - \frac{1}{N} E\left[\mathbf{q}^H \sum_{i_1=1}^{U} \sum_{m_1=1}^{M_i} \mathbf{V}_{i_1,m_1} \sum_{i_2=1}^{U} \sum_{m_2=1}^{M_i} \mathbf{V}_{i_2,m_2}^H \mathbf{q}\right]. \end{aligned} \tag{A.1}$$

It is straightforward to show that (the definition of β_i is given in (33))

$$\mathbf{q}^H \mathbf{V}_{i,m} = (\underline{\mathrm{c}} \otimes \underline{\mathrm{h}})^H (V_{i,m} \otimes \underline{\mathrm{a}}_i) = \left(\underline{\mathrm{c}}^T V_{i,m}\right) \otimes \left(\underline{\mathrm{h}}^H \underline{\mathrm{a}}_i\right) = N\beta_i \sum_{l=0}^{L-1} V_{i,m}(l) c(l). \tag{A.2}$$

Using the orthogonal property of the spreading sequence A1), (A.1) becomes

$$\begin{aligned} E[y_1(n)] &= LN - NE\left[\sum_{i_1=1}^{U} \beta_{i_1} \sum_{m_1=1}^{M_{i_1}} \sum_{l_1=0}^{L-1} V_{i_1,m_1}(l_1) c(l_1) \sum_{i_2=1}^{U} \beta_{i_2}^* \sum_{m_2=1}^{M_{i_2}} \sum_{l_2=0}^{L-1} V_{i_2,m_2}^*(l_2) c(l_2)\right] \\ &= LN - N \sum_{i=1}^{U} |\beta_i|^2 \sum_{m=1}^{M_i} \sum_{l=0}^{L-1} |V_{i,m}(l)|^2 c^2(l) \\ &= N\left(L - \sum_{i=1}^{U} M_i |\beta_i|^2\right). \end{aligned} \tag{A.3}$$

Due to the zero-mean property of noise (assumption A2), $E[y_2(n)] = 0$. Accordingly,

$$E[y(n)] = E[y_1(n)] = N\left(L - \sum_{i=1}^{U} M_i |\beta_i|^2\right). \tag{A.4}$$

It is clear from (A.4) that the increase in the space dimensionality from L to NL does not simply translate into a corresponding increase in the desired mean value, or subsequently in the processing gain. Also, from assumption A3), the cross-correlation between $y_1(n)$ and $y_2(n)$ is zero, i.e.,

$$E[y_1^*(n) y_2(n)] = E[y_1(n) y_2^*(n)] = 0. \tag{A.5}$$

Therefore, the mean square value of the decision variable is made up of only two terms,

$$E\left[|y(n)|^2\right] = E\left[|y_1(n)|^2\right] + E\left[|y_2(n)|^2\right]. \tag{A.6}$$

The first term is the mean square value of $y_1(n)$. From (26), we have

$$\begin{aligned}
E\left[|y_1(n)|^2\right] &= E\left[\mathbf{q}^H\mathbf{P}\mathbf{q}\mathbf{q}^H\mathbf{P}^H\mathbf{q}\right] \\
&= E\left[\mathbf{q}^H\left(\mathbf{I}_N - \frac{1}{N}\mathbf{V}\mathbf{V}^H\right)\mathbf{q}\mathbf{q}^H\left(\mathbf{I}_{LN} - \frac{1}{LN}\mathbf{V}\mathbf{V}^H\right)\mathbf{q}\right] \\
&= E\left[\mathbf{q}^H\mathbf{q}\mathbf{q}^H\mathbf{q}\right] - \frac{2}{N}E\left[\mathbf{q}^H\mathbf{q}\mathbf{q}^H\mathbf{V}\mathbf{V}^H\mathbf{q}\right] + \frac{1}{N^2}E\left[\mathbf{q}^H\mathbf{V}\mathbf{V}^H\mathbf{q}\mathbf{q}^H\mathbf{V}\mathbf{V}^H\mathbf{q}\right] \\
&= (LN)^2 - 2LN^2\sum_{i=1}^{U} M_i|\beta_i|^2 \\
&\quad + N^2E\left[\sum_{i_1=1}^{U}\beta_{i_1}\sum_{m_1=1}^{M_{i_1}}\sum_{l_1=0}^{L-1}V_{i_1,m_1}(l_1)c(l_1)\sum_{i_2=1}^{U}\beta_{i_2}^*\sum_{m_2=1}^{M_{i_2}}\sum_{l_2=0}^{L-1}V_{i_2,m_2}^*(l_2)c(l_2)\right. \\
&\qquad \left.\times\sum_{i_3=1}^{U}\beta_{i_3}\sum_{m_3=1}^{M_{i_3}}\sum_{l_3=0}^{L-1}V_{i_3,m_3}(l_3)c(l_3)\sum_{i_4=1}^{U}\beta_{i_4}^*\sum_{m_4=1}^{M_{i_4}}\sum_{l_4=0}^{L-1}V_{i_4,m_4}^*(l_4)c(l_4)\right] \\
&= (LN)^2 - 2LN^2\sum_{i=1}^{U} M_i|\beta_i|^2 \\
&\quad + N^2E\left[\sum_{i_1=i_2=1}^{U}\beta_{i_1}\beta_{i_2}^*\sum_{m_1=m_2=1}^{M_{i_1}}\sum_{l_1=l_2=0}^{L-1}V_{i_1,m_1}(l_1)V_{i_2,m_2}^*(l_2)c(l_1)c(l_2)\right. \\
&\qquad \left.\times\sum_{i_3=i_4=1}^{U}\beta_{i_3}\beta_{i_4}^*\sum_{m_3=m_4=1}^{M_{i_3}}\sum_{l_3=l_4=0}^{L-1}V_{i_3,m_3}(l_3)V_{i_4,m_4}^*(l_4)c(l_3)c(l_4)\right] \\
&\quad + N^2E\left[\sum_{i_1=i_4=1}^{U}\beta_{i_1}\beta_{i_4}^*\sum_{m_1=m_4=1}^{M_{i_1}}\sum_{l_1=l_4=0}^{L-1}V_{i_1,m_1}(l_1)V_{i_4,m_4}^*(l_4)c(l_1)c(l_4)\right. \\
&\qquad \left.\times\sum_{i_3=i_2=1}^{U}\beta_{i_3}\beta_{i_2}^*\sum_{m_3=m_2=1}^{M_{i_3}}\sum_{l_3=l_2=0}^{L-1}V_{i_3,m_3}(l_3)V_{i_2,m_2}^*(l_2)c(l_3)c(l_2)\right] \\
&\quad + N^2E\left[\sum_{i_1=i_3=1}^{U}\beta_{i_1}\beta_{i_3}\sum_{m_1=1}^{M_{i_1}}\sum_{m_3=1}^{M_{i_3}}\sum_{l_1=l_3=0}^{L-1}V_{i_1,m_1}(l_1)V_{i_3,m_3}(l_3)c(l_1)c(l_3)\right. \\
&\qquad \left.\times\sum_{i_2=i_4=1}^{U}\beta_{i_2}^*\beta_{i_4}^*\sum_{m_2=1}^{M_{i_2}}\sum_{m_4=1}^{M_{i_4}}\sum_{l_2=l_4=0}^{L-1}V_{i_2,m_2}^*(l_2)V_{i_4,m_4}^*(l_4)c(l_2)c(l_4)\right] \\
&\quad - 2N^2E\left[\sum_{i_1=i_2=i_3=i_4=1}^{U}\beta_{i_1}\beta_{i_2}^*\beta_{i_3}\beta_{i_4}^*\sum_{m_1=m_2=m_3=m_4=1}^{M_{i_1}}\right. \\
&\qquad \left.\times\sum_{l_1=l_2=l_3=l_4=0}^{L-1}V_{i_1,m_1}(l_1)V_{i_2,m_2}^*(l_2)V_{i_3,m_3}(l_3)V_{i_4,m_4}^*(l_4)c(l_1)c(l_2)c(l_3)c(l_4)\right]
\end{aligned}$$

$$= (LN)^2 - 2LN^2 \sum_{i=1}^{U} M_i|\beta_i|^2 + N^2 \left(2 \left(\sum_{i=1}^{U} M_i|\beta_i|^2 \right)^2 + \left| \sum_{i=1}^{U} \beta_i^2 \gamma_i \right|^2 - 2 \sum_{i=1}^{U} \xi_i |\beta_i|^4 \right), \quad \text{(A.7)}$$

where

$$\gamma_i = \sum_{m_1=1}^{M_i} \sum_{m_2=1}^{M_i} \sum_{l=1}^{L-1} V_{i,m_1}(l) V_{i,m_2}(l) \quad \text{(A.8)}$$

and

$$\xi_i = \sum_{m=1}^{M_i} \sum_{l=1}^{L-1} |V_{i,m}(l)|^4 . \quad \text{(A.9)}$$

In practice, γ_i takes negligible values, and equation (A.7) can be simplified to

$$E\left[|y_1(n)|^2\right] = (LN)^2 - 2LN^2 \sum_{i=1}^{U} M_i|\beta_i|^2 + N^2 \left(2 \left(\sum_{i=1}^{U} M_i|\beta_i|^2 \right)^2 - 2 \sum_{i=1}^{U} \xi_i |\beta_i|^4 \right). \quad \text{(A.10)}$$

The value of ξ_i depends on the type of interference signals. Specifically, when the ith interference signal is made up of a single FM or a number of uncorrelated FM signal components, then the basis vectors are of constant modulus, and

$$\xi_i = \frac{M_i}{L}. \quad \text{(A.11)}$$

The second term of (A.6) is the mean-square value of $y_2(n)$,

$$\begin{aligned} E\left[|y_2(n)|^2\right] &= E\left[\mathbf{q}^H \mathbf{P} \mathbf{B}(k) \mathbf{B}^H(k) \mathbf{P}^H \mathbf{q}\right] \\ &= \sigma E\left[\mathbf{q}^H \mathbf{P} \mathbf{P}^H \mathbf{q}\right] = \sigma E\left[\mathbf{q}^H \mathbf{P} \mathbf{q}\right] = \sigma N \left(L - \sum_{i=1}^{U} M_i|\beta_i|^2 \right). \end{aligned} \quad \text{(A.12)}$$

The variance of $y(n)$ is given by

$$\begin{aligned} var[y(n)] &= E[|y(n)|^2] - E^2[y(n)] \\ &= E[|y_1(n)|^2] + E[|y_2(n)|^2] - E^2[y_1(n)] \\ &= (LN)^2 - 2LN^2 \sum_{i=1}^{U} M_i|\beta_i|^2 + N^2 \left(2 \left(\sum_{i=1}^{U} M_i|\beta_i|^2 \right)^2 - \frac{1}{L} \sum_{i=1}^{U} M_i|\beta_i|^4 \right) \\ &\quad + \sigma N \left(L - \sum_{i=1}^{U} M_i|\beta_i|^2 \right) - N^2 \left(L - \sum_{i=1}^{U} M_i|\beta_i|^2 \right)^2 \\ &= N^2 \left(\left(\sum_{i=1}^{U} M_i|\beta_i|^2 \right)^2 - 2 \sum_{i=1}^{U} \xi_i |\beta_i|^4 \right) + \sigma N \left(L - \sum_{i=1}^{U} M_i|\beta_i|^2 \right). \end{aligned} \quad \text{(A.13)}$$

Equation (32) follows by using the results of (A.4) and (A.13).

References

[1] H. V. Poor and L. A. Rusch, "Narrowband interference suppression in spread-spectrum CDMA," *IEEE Personal Comm. Mag.*, vol. 1, no. 8, pp. 14–27, Aug. 1994.

[2] J. D. Laster and J. H. Reed, "Interference rejection in digital wireless communications," *IEEE Signal Processing Mag.*, vol. 14, no. 3, pp. 37–62, May 1997.

[3] L. B. Milstein, "Interference rejection techniques in spread spectrum communications," *Proc. IEEE*, vol. 76, no. 6, pp. 657–671, June 1988.

[4] J. Wang and L. B. Milstein, "Adaptive LMS filters for cellular CDMA overlay," *IEEE J. Select. Areas Commun.*, vol. 14, no. 8, pp. 1548–1559, Oct. 1996.

[5] S. Sandberg, "Adapted demodulation for spread-spectrum receivers which employ transform-domain interference excision," *IEEE Trans. Commun.*, vol. 43, pp. 2502–2510, Sept. 1995.

[6] L. A. Rusch and H. Poor, "Multiuser detection techniques for narrow-band interference suppression in spread spectrum communications," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 1725–1737, Feb./Mar./Apr. 1995.

[7] H. Fathallah and L. A. Rusch, "A subspace approach to adaptive narrow-band interference suppression in DSSS," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1575–1585, Dec. 1997.

[8] M. Lops, G. Ricci, and A. T. Tulino, "Narrow-band-interference suppression in multiuser CDMA systems," *IEEE Trans. Signal Processing*, vol. 46, no. 9, pp. 1163–1175, Sept. 1998.

[9] L. A. Rusch, "MMSE detector for narrow-band interference suppression in DS spread spectrum," in *Proc. Interference Rejection and Signal Separation in Wireless Commun. Symp.*, Newark, NJ, March 1996.

[10] M. G. Amin and A. Akansu, "Time-frequency for interference excision in spread-spectrum communications," in "Highlights of signal processing for communications:

celebrating a half century of signal processing," *IEEE Signal Processing Mag.*, vol. 16, no. 2, March 1999.

[11] M. G. Amin, "Interference mitigation in spread spectrum communication systems using time-frequency distribution," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp.90–102, Jan. 1997.

[12] M. G. Amin, C. Wang, and A. Lindsey, "Optimum interference excision in spread spectrum communications using open loop adaptive filters," *IEEE Trans. Signal Processing*, vol. 47, no. 7, pp.1966–1976, July 1999.

[13] B. K. Poh, T. S. Quek, C. M. S. See, and A. C. Kot, "Suppression of strong narrowband interference using eigen-structure-based algorithm," in *Proc. Milcom*, pp. 1205–1208, July 1995.

[14] A. Haimovich and A. Vadhri, "Rejection of narrowband interferences in PN spread spectrum systems using an eigenanalysis approach," in *Proc. IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, Quebec, Canada, pp. 1002–1006, June 1994.

[15] F. Hlawatsch and W. Kozek, "Time-frequency projection filters and time-frequency signal expansions," *IEEE Trans. Signal Processing*, vol. 42, no. 12, pp. 3321–3334, Dec. 1994.

[16] M. G. Amin and G. R. Mandapati, "Nonstationary interference excision in spread spectrum communications using projection filtering methods," in *Proc. 32nd Annual Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1998.

[17] S. Barbarossa and A. Scaglione, "Adaptive time-varying cancellation of wideband interferences in spread-spectrum communications based on time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 47, no. 4, pp. 957–965, April 1999.

[18] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal," *Proc. IEEE*, vol. 80, no. 12, Dec. 1990.

[19] P. Loughlin and K. Davidson, "Instantaneous bandwidth of multicomponent signals,"

in *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, vol. 3807, pp. 546–551, July 1999.

[20] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, 1995.

[21] P. Rao and F. J. Taylor, "Estimation of the instantaneous frequency using the discrete Wigner distribution," *Electronics Lett.*, vol. 26, no. 4, pp. 246–248, Feb. 1990.

[22] M. Wang, A. Chan, and C. Chui, "Linear frequency-modulated signal detection using Radon-ambiguity transform," *IEEE Trans. Signal Processing*, vol. 46, no. 3, pp. 571–586, March 1998.

[23] P. Shan and A. A. Beex, "FM interference suppression in spread spectrum communications using time-varying autoregressive model based instantaneous frequency estimation," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, pp. 2559–2562, March 1999.

[24] R. S. Ramineni, M. G. Amin, and A. R. Lindsey, "Performance analysis of subspace projection techniques for interference excision in DSSS communications," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Istanbul, Turkey, June 2000.

[25] W. Mu, Y. Zhang, and M. G. Amin, "Bilinear signal synthesis in array processing," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Salt Lake City, UT, May 2001.

[26] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. 34, no. 3, pp. 276–280, March 1986.

[27] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, no. 10, pp. 1553–1560, Oct. 1988.

[28] A. Belouchrani and M. Amin, "Time-frequency MUSIC," *IEEE Signal Processing Lett.*, vol. 6, no. 5, pp. 109–110, May 1999.

[29] Y. Zhang, W. Mu, and M. G. Amin, "Time-frequency maximum likelihood methods for direction finding," *J. Franklin Inst.*, vol. 337, no. 4, pp. 483–497, July 2000.

[30] Y. Zhang, W. Mu, and M. G. Amin, "Subspace analysis of spatial time-frequency distribution matrices," *IEEE Trans. Signal Processing*, vol. 49, no. 4, pp. 747–759, Apr. 2001.

[31] J. F. Cardoso, A. Belouchrani, K. Abed Maraim, and E. Moulines, "A blind source separation technique using second order statistics," *IEEE Trans. Signal Processing*, vol. 45, no. 2, pp. 434–444, Feb. 1997.

[32] A. Belouchrani and M. Amin, "Blind source separation based on time-frequency signal representation," *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. 2888–2898, Nov. 1998.

[33] Y. Zhang and M. G. Amin, "Blind separation of sources based on their time-frequency signatures," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Istanbul, Turkey, June 2000.

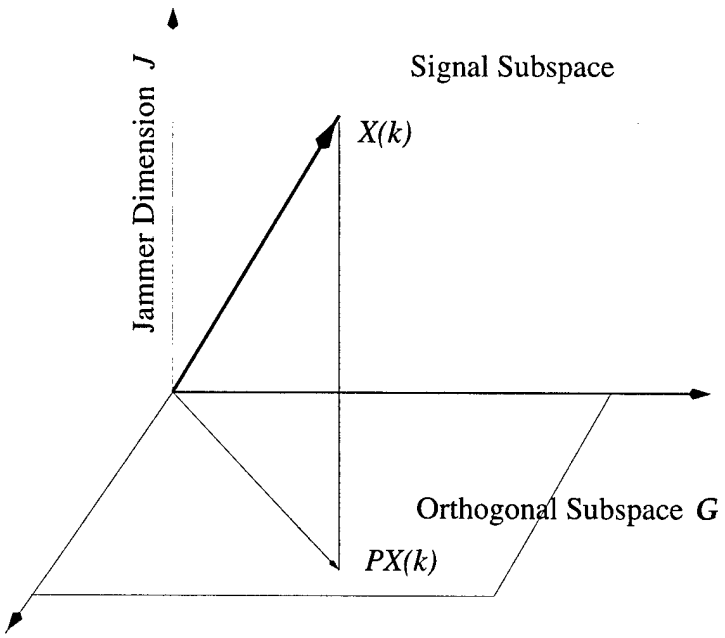


Fig. 1 Jammer suppression by subspace projection.

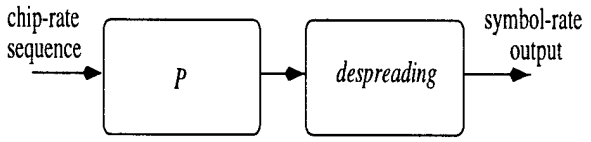


Fig. 2 Block diagram of single-sensor subspace projection.

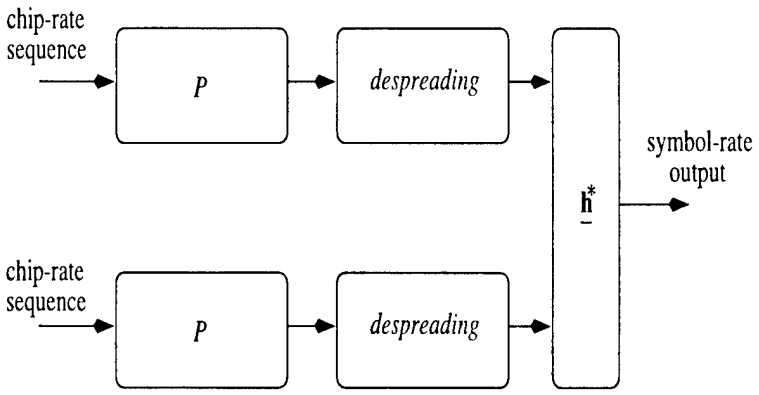


Fig. 3 Block diagram of independent multi-sensor subspace projection.

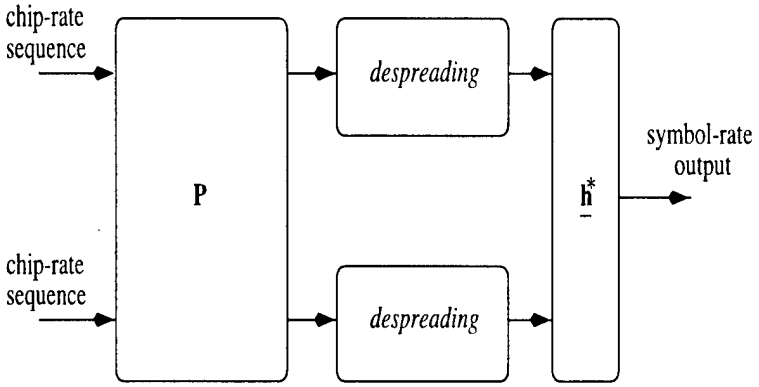


Fig. 4 Block diagram of proposed multi-sensor subspace projection.

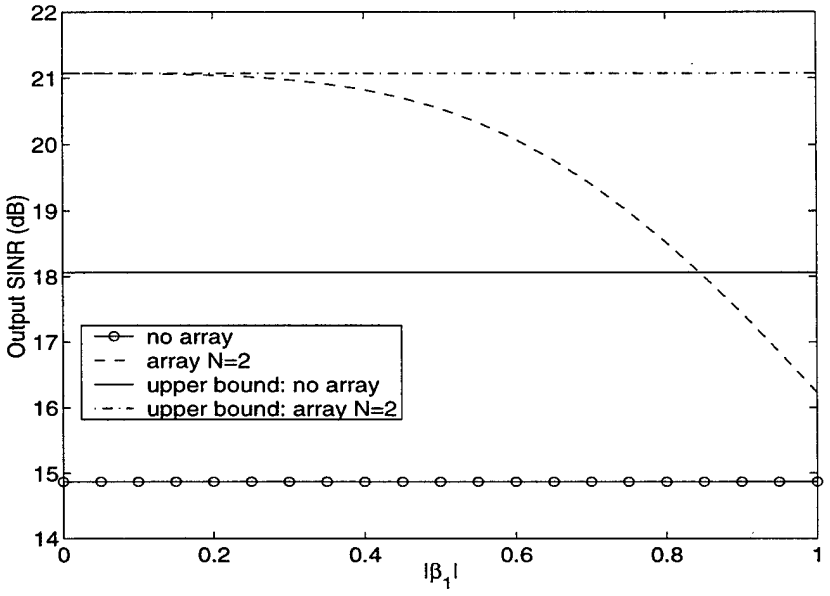


Fig. 5 Output SINR versus $|\beta_1|$ (input SNR=0dB, L=64, U=1, M=7).

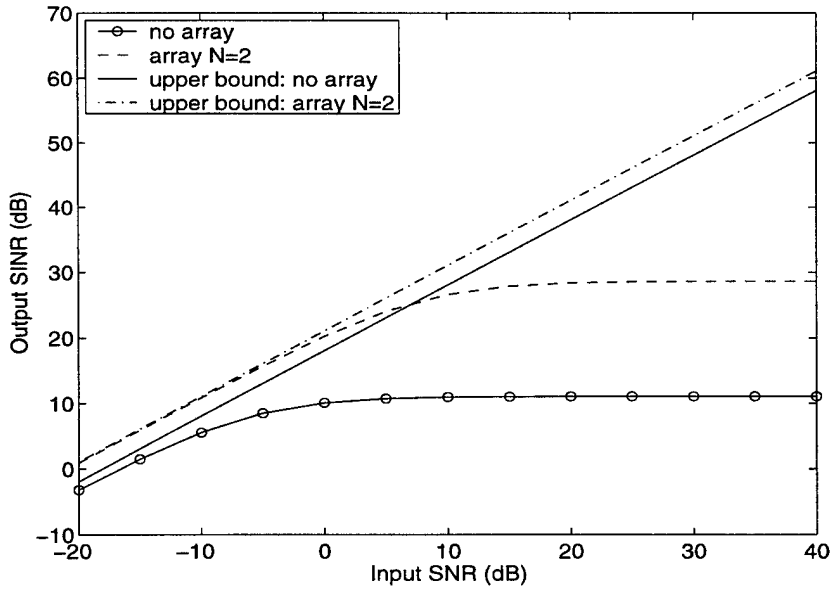


Fig. 6 Output SINR versus input SNR
(L=64, U=2, $M_1 = M_2 = 7$, $\theta_D = 0°$, $\theta_J = [40°, 60°]$).

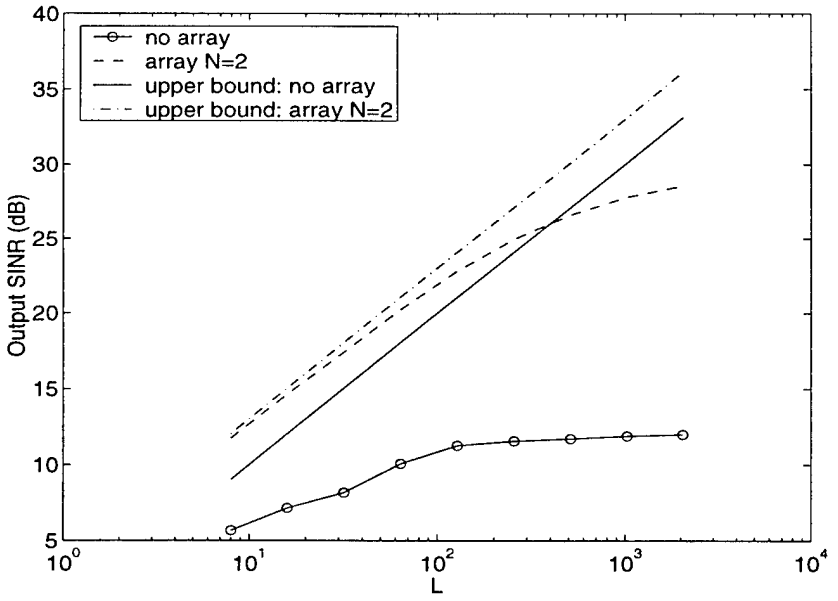


Fig. 7 Output SINR versus the number of chips per symbol (L) (input SNR=0dB, U=2, $M_1 = M_2 = 7$, $\theta_D = 0^o$, $\theta_J = [40^o, 60^o]$).

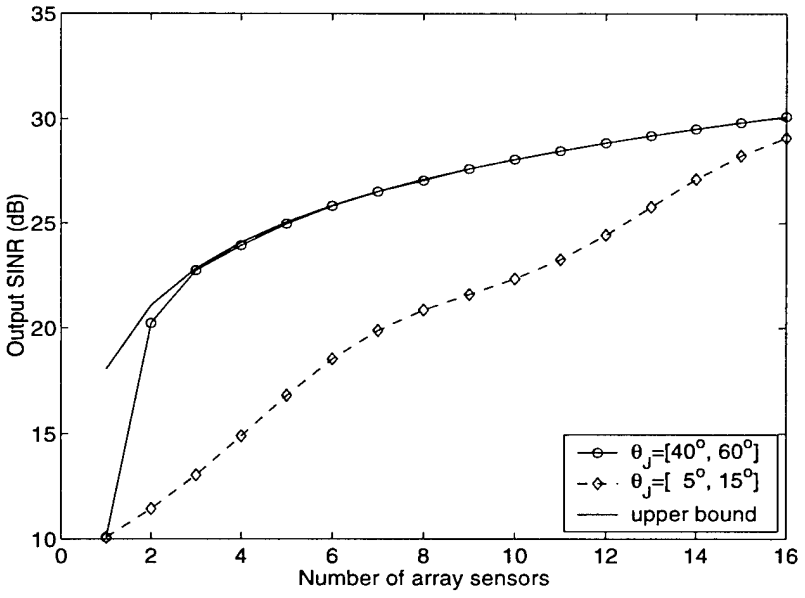


Fig. 8 Output SINR versus the number of array sensors (input SNR=0dB, L=64, U=2, $M_1 = M_2 = 7$).

Estimating parameters of multiple wideband polynomial-phase sources in sensor arrays

A.B. Gershman[†], M. Pesavento[†], and Moeness G. Amin[‡]
[†] Department of Electrical and Computer
McMaster University, Hamilton, Ontario, L8S 4K1, Canada
[‡] Department of Electrical and Computer Engineering,
Villanova University, Villanova, PA 19085

Abstract

In this paper, we consider the problem of estimating the parameters of multiple wideband polynomial-phase signal (PPS) sources in sensor arrays. A new maximum likelihood (ML) direction of arrival (DOA) estimator is introduced and the exact Cramér-Rao bound (CRB) is derived for the general case of multiple constant-amplitude polynomial-phase sources. Since the proposed exact ML estimator is computationally intensive, an approximate solution is proposed, originating from the analysis of the log-likelihood (LL) function in the single chirp signal case. As a result, a new form of spatio-temporal matched filter (referred to as the *chirp beamformer*) is derived, which is applicable to "well-separated" sources that have distinct time-frequency or/and spatial signatures. This beamforming approach requires solving a 3D optimization problem and, therefore, enjoys essentially simpler implementation than that entailed by the exact ML. Simulation results are presented, illustrating the performance of the estimators and validating our theoretical CRB analysis.

This work is supported in parts by the Natural Sciences and Engineering Research Council (NSERC) of Canada under Grant Ministry of Energy, Science, and Communications the Air Force Research Lab under contract F30602-00-0515, and the Office of Naval Research (ONR) under Grant N00014-98-1-1076.

I. Introduction

Estimating the parameters of polynomial-phase signals (PPS's) is an important problem because linear and nonlinear frequency-modulated (FM) signals are encountered in several practical applications. For example, in synthetic aperture radar (SAR), synthetic aperture sonar (SAS), inverse SAR and SAS (ISAR and ISAS), and in Doppler radar and sonar imaging, the returns are FM signals [1]-[4]. Furthermore, FM signal waveforms can be *intentionally* transmitted in multi-sensor systems. For example, chirp signal waveforms are widely used for pulse-compression in radar and sonar [1]. Both cases of constant-amplitude and time-varying amplitude FM signals have attracted much attention in the literature [3]-[8]. Recently, there has been a growing interest in estimating the parameters of multiple PPS sources impinging on a multi-sensor array [9]-[14]. Several authors have solved this problem using narrowband assumptions, where the signal spatial signatures are assumed to remain invariant over the observation period. In particular, a new spatial time-frequency distribution (STFD) concept has been developed and employed for direction finding of narrowband FM sources using subspace techniques [10]-[11]. Several exact and approximate ML algorithms for this estimation problem have been proposed [9]. Promising extensions of the above-mentioned narrowband approaches to the case of wideband FM signals have been recently reported [12]-[15]. However, these wideband methods suffer from a high computational cost and restrictive assumptions. In particular, the application of the wideband STFD approach [13]-[14] is severely restricted by short sliding data window lengths, whereas the consideration in [12] is limited by the assumption of linear FM signals with the central frequency which is known and identical for each source. The iterative approach reported in [15] may lead to strongly biased DOA estimates [13], and its convergence is not guaranteed.

It is also important to stress that the conventional wideband techniques (such as coherent-subspace MUSIC [16]-[17] and wideband ML [18]) have only a limited application to the

PPS case because they do not take any advantage of the specific PPS structure. Also, the wideband coherent-subspace-based techniques require preliminary information about the source localization angular sectors or initial DOA estimates to compute focussing matrices. The focussing operation itself can be done only approximately and this may lead to a strong bias in the resulting DOA estimates. Furthermore, in the PPS case these subspace-based techniques are severely restricted by the window length (this restriction is similar to that of the wideband STFD approach [13]) and their performance may be very critical to the choice of several free parameters, such as the focussing frequency, the width of array interpolation sectors, etc.

In this paper, we derive a new form of the ML estimator of the parameters of multiple wideband constant-amplitude PPS sources received by a sensor array. Our technique is free of the above-mentioned restrictions on the signal waveform parameters and the length of the observation interval. Novel explicit expressions for the corresponding CRB on the accuracy of estimating the model parameters are derived.

Although the presented ML estimator concentrates the problem at hand with respect to the signal nuisance parameters (initial amplitudes), its computational cost may be still very high, since it involves a nonlinear optimization over the parameter space of a high dimension. Since the proposed exact ML estimator is computationally intensive, an approximate solution is considered, originating from the analysis of the LL function in the single linear FM source case. Using this approximation, we derive a new form of spatio-temporal matched filter (referred to as the *chirp beamformer*), which is applicable to the wide class of scenarios with "well-separated" (either in DOA or in any of frequency parameters) sources. The developed chirp beamforming approach entails solving a 3D optimization problem and, therefore, enjoys essentially simpler implementation than the exact ML technique.

Our paper is organized as follows. The array signal model used in the underlying problem is presented in Section 2. Based on this model, the ML estimator is derived in Section

3, and the exact CRB on the accuracy of estimating the model parameters is established in Section 4. In Section 5, we present the chirp beamformer as a simplified alternative for the ML estimator. Simulation results are presented in Section 6. They illustrate the performance of the estimators and validate our CRB analysis.

II. Array Signal Model

Assume that L wideband constant-amplitude PPS's impinge on a linear array of M omnidirectional sensors. Then, the vector array outputs obey the following model

$$\boldsymbol{x}(t) = \boldsymbol{A}(t)\boldsymbol{s}(t) + \boldsymbol{n}(t)\,, \quad t = 0, 1, \ldots, N-1 \tag{1}$$

where $\boldsymbol{A}(t)$ is the $M \times L$ time-varying direction matrix, $\boldsymbol{s}(t)$ is the $L \times 1$ vector of wideband nonstationary source waveforms, $\boldsymbol{n}(t)$ is the $M \times 1$ vector of complex circularly Gaussian zero-mean temporally and spatially white sensor noise, and N is the number of snapshots.

The lth polynomial-phase source waveform can be modeled as

$$s_l(t) = \alpha_l e^{j(\omega_{l,0}t + \omega_{l,1}t^2/2 + \cdots + \omega_{l,K-1}t^K/K)} = \alpha_l g(\boldsymbol{\omega}_l, t) \tag{2}$$

where

$$g(\boldsymbol{\omega}_l, t) \triangleq \exp\left\{ j \sum_{k=0}^{K-1} \omega_{l,k} \frac{t^{k+1}}{k+1} \right\} \tag{3}$$

α_l is the initial complex amplitude, $\omega_{l,k}$ $(l = 1, 2, \ldots, L;\ k = 0, 1, \ldots, K-1)$ are the unknown discrete-time frequency parameters, and

$$\tilde{\omega}_l(t) = \sum_{k=0}^{K-1} \omega_{l,k} t^k \tag{4}$$

is the discrete-time instantaneous frequency of the lth waveform. It is important to stress that the discrete- and continuous-time instantaneous signal frequencies are related as

$$\tilde{\omega}_l(t) = \Delta t\, \tilde{\omega}_l^{\mathrm{ct}}(t) \tag{5}$$

where Δt is the sampling interval and $\tilde{\omega}_l^{\mathrm{ct}}(t)$ is the continuous-time (physical) instantaneous frequency of the lth PPS. The $K \times 1$ vector

$$\boldsymbol{\omega}_l \triangleq [\omega_{l,0}, \omega_{l,1}, \ldots, \omega_{l,K-1}]^T \tag{6}$$

contains the unknown discrete-time frequency parameters of the lth signal, and K is the order of the polynomial-phase model[1].

The direction matrix

$$\begin{aligned} \boldsymbol{A}(\boldsymbol{\theta}, \tilde{\boldsymbol{\omega}}(t)) &= [\boldsymbol{a}(\theta_1, \tilde{\omega}_1(t)), \boldsymbol{a}(\theta_2, \tilde{\omega}_2(t)), \ldots, \boldsymbol{a}(\theta_L, \tilde{\omega}_L(t))] \\ &= [\boldsymbol{a}(\theta_1, \boldsymbol{\omega}_1, t), \boldsymbol{a}(\theta_2, \boldsymbol{\omega}_2, t), \ldots, \boldsymbol{a}(\theta_L, \boldsymbol{\omega}_L, t)] \\ &= \boldsymbol{A}(\boldsymbol{\theta}, \boldsymbol{\omega}, t) \end{aligned} \tag{7}$$

consists of the time-varying steering vectors

$$\boldsymbol{a}(\theta_l, \boldsymbol{\omega}_l, t) = \left[1, \exp\left\{j \frac{\tilde{\omega}_l(t)}{c\,\Delta t} d_1 \sin\theta_l\right\}, \ldots, \exp\left\{j \frac{\tilde{\omega}_l(t)}{c\,\Delta t} d_{M-1} \sin\theta_l\right\}\right]^T \tag{8}$$

where

$$\boldsymbol{\theta} \triangleq [\theta_1, \theta_2, \ldots, \theta_L]^T \tag{9}$$

$$\boldsymbol{\omega} \triangleq \left[\boldsymbol{\omega}_1^T, \boldsymbol{\omega}_2^T, \ldots, \boldsymbol{\omega}_L^T\right]^T \tag{10}$$

$$\tilde{\boldsymbol{\omega}}(t) \triangleq [\tilde{\omega}_1(t), \tilde{\omega}_2(t), \ldots, \tilde{\omega}_L(t)]^T \tag{11}$$

θ_l is the DOA of the lth source, and d_i is the spacing between the first and the $(i+1)$th array sensors. We stress that the relationship (5) is used in (8) to express the steering vector as a function of the discrete-time frequency parameters. As follows from (7), the direction matrix can be written as a function either of the frequency parameters (10) or of instantaneous frequencies (11). Note that in (8) it is assumed that the instantaneous signal frequencies $\tilde{\omega}_l(t)$ $(t = 1, \ldots, L)$ do not change during the time necessary for a wave

[1]In what follows, the order K is assumed to be known.

to travel across the array aperture[2].

Using (2)-(11), model (1) can be rewritten as

$$\begin{aligned} \boldsymbol{x}(t) &= \boldsymbol{A}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{G}(\boldsymbol{\omega},t)\boldsymbol{\alpha}+\boldsymbol{n}(t) \\ &= \tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{\alpha}+\boldsymbol{n}(t) \end{aligned} \tag{12}$$

where

$$\boldsymbol{\alpha} \triangleq [\alpha_1,\ldots,\alpha_L]^T \tag{13}$$

$$\boldsymbol{G}(\boldsymbol{\omega},t) \triangleq \mathrm{diag}\{g(\boldsymbol{\omega}_1,t),\ldots,g(\boldsymbol{\omega}_L,t)\} \tag{14}$$

$$\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t) \triangleq \boldsymbol{A}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{G}(\boldsymbol{\omega},t) \tag{15}$$

Note that all nuisance parameters (the initial source amplitudes) are now included in the vector $\boldsymbol{\alpha}$.

III. Maximum Likelihood Estimator

In this section, we derive a novel ML estimator of the source DOA's and frequency parameters based on the assumption that the initial signal amplitudes are constant (deterministic) values. The negative log-likelihood (LL) function is given by

$$\begin{aligned} \mathcal{L}_N(\Theta) &= \sum_{t=0}^{N-1} ||\boldsymbol{x}(t)-\boldsymbol{A}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{G}(\boldsymbol{\omega},t)\boldsymbol{\alpha}||^2 \\ &= \sum_{t=0}^{N-1} ||\boldsymbol{x}(t)-\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{\alpha}||^2 \end{aligned} \tag{16}$$

where the $(LK+2L)\times 1$ vector of unknown model parameters is defined as

$$\Theta \triangleq \left[\boldsymbol{\theta}^T,\boldsymbol{\omega}^T,\boldsymbol{\alpha}^T\right]^T$$

[2]This assumption means that the signals remain narrowband in each snapshot, i.e. their instantaneous bandwidths are small compared with the inverse of the wavefront propagation time across the array. However, the signals are assumed to be wideband at the full observation interval of N samples because the propagation time across the aperture is usually much smaller than the sampling interval.

Rewrite (16) as

$$\begin{aligned}\mathcal{L}_N(\Theta) &= \boldsymbol{\alpha}^H \left\{ \sum_{t=0}^{N-1} \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t) \right\} \boldsymbol{\alpha} - \left\{ \sum_{t=0}^{N-1} \boldsymbol{x}^H(t)\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t) \right\} \boldsymbol{\alpha} \\ &- \boldsymbol{\alpha}^H \left\{ \sum_{t=0}^{N-1} \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{x}(t) \right\} + \sum_{t=0}^{N-1} \boldsymbol{x}^H(t)\boldsymbol{x}(t) \end{aligned} \tag{17}$$

The minimization of $\mathcal{L}_N$ over $\boldsymbol{\alpha}$ yields

$$\hat{\boldsymbol{\alpha}} = \left\{ \sum_{t=0}^{N-1} \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t) \right\}^{-1} \left\{ \sum_{t=0}^{N-1} \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{x}(t) \right\} \tag{18}$$

Substituting (18) into (17), we obtain the concentrated negative LL function

$$\begin{aligned}\mathcal{L}_N(\boldsymbol{\theta},\boldsymbol{\omega}) &= \sum_{t=0}^{N-1} \boldsymbol{x}^H(t)\boldsymbol{x}(t) \\ &- \left\{ \sum_{t=0}^{N-1} \boldsymbol{x}^H(t)\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t) \right\} \left\{ \sum_{t=0}^{N-1} \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t) \right\}^{-1} \\ &\times \left\{ \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{x}(t) \right\} \end{aligned} \tag{19}$$

Ignoring the constant terms, the positive concentrated LL function is given by

$$\begin{aligned}\mathcal{L}_P(\boldsymbol{\theta},\boldsymbol{\omega}) &= \left\{ \sum_{t=0}^{N-1} \boldsymbol{x}^H(t)\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t) \right\} \left\{ \sum_{t=0}^{N-1} \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t) \right\}^{-1} \\ &\times \left\{ \sum_{t=0}^{N-1} \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{x}(t) \right\} \end{aligned} \tag{20}$$

The ML estimator is

$$[\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\omega}}] = \arg \max_{\boldsymbol{\theta},\boldsymbol{\omega}} \mathcal{L}_P(\boldsymbol{\theta},\boldsymbol{\omega}) \tag{21}$$

The above estimator jointly estimates the source directions and their frequency parameters $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$, respectively, and generally requires a highly nonlinear optimization of (20) over these variables. However, if properly initialized, the optimization of the LL function may be implemented by means of simple local optimization algorithms.

IV. Cramér-Rao Bound

In this section, we derive novel explicit expressions for the exact CRB on the accuracy of estimating the signal model parameters.

The observations (12) satisfy the following statistical model:

$$\boldsymbol{x}(t) \sim \mathcal{N}_C\left(\tilde{\boldsymbol{A}}(\boldsymbol{\theta}, \boldsymbol{\omega}, t)\boldsymbol{\alpha}, \sigma^2 \boldsymbol{I}\right) \tag{22}$$

where $\mathcal{N}_C$ denotes the complex multivariate circularly Gaussian probability density function, and σ^2 is the noise variance.

For mathematical convenience, we redefine the signal frequency parameters in an order different from (6) and (10). Let $\boldsymbol{\Omega}$ denote the matrix

$$\boldsymbol{\Omega} \triangleq \begin{bmatrix} \omega_{1,0} & \omega_{2,0} & \cdots & \omega_{L,0} \\ \omega_{1,1} & \omega_{2,1} & \cdots & \omega_{L,1} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{1,K-1} & \omega_{2,K-1} & \cdots & \omega_{L,K-1} \end{bmatrix} \tag{23}$$

Using (23), we can alternatively rewrite equation (10) as

$$\boldsymbol{\omega} \triangleq \text{vec}\left\{\boldsymbol{\Omega}\right\} \tag{24}$$

where vec$\{\cdot\}$ represents the so-called vectorization operator stacking the columns of a matrix to form a column vector. In what follows, we will also use another definition of the vector of the signal parameters,

$$\begin{aligned} \boldsymbol{\nu} &\triangleq \text{vec}\left\{\boldsymbol{\Omega}^T\right\} \\ &= \left[\boldsymbol{\nu}_0^T, \ldots, \boldsymbol{\nu}_{K-1}^T\right]^T \end{aligned} \tag{25}$$

where

$$\boldsymbol{\nu}_k \triangleq \left[\omega_{1,k}, \omega_{2,k}, \ldots, \omega_{L,k}\right]^T \tag{26}$$

The two alternative definitions (24) and (25) are equivalent, i.e., in principle either one can be applied in the underlying problem. However, the use of (25) in lieu of (24) in the

definition of the full parameter vector helps simplifying the derivations, as it corresponds to a more convenient order of the subblocks of the Fisher Information Matrix (FIM). We note that the first definition (24) will be used herein to denote the frequency parameters as the arguments of matrices or scalar functions, whereas the second definition (25) will be only exploited to provide a proper order of the FIM elements. Using (25), the $(LK+3L+1)\times 1$ vector of unknown real parameters can be defined as

$$\begin{aligned} \boldsymbol{\Psi} &\triangleq \left[\boldsymbol{\theta}^T, \mathrm{Re}\{\boldsymbol{\alpha}\}^T, \mathrm{Im}\{\boldsymbol{\alpha}\}^T, \boldsymbol{\nu}^T, \sigma^2\right]^T \\ &= \left[\boldsymbol{\theta}^T, \breve{\boldsymbol{\alpha}}^T, \boldsymbol{\nu}^T, \sigma^2\right]^T \end{aligned} \tag{27}$$

where

$$\breve{\boldsymbol{\alpha}} = \left[\mathrm{Re}\{\boldsymbol{\alpha}\}^T, \mathrm{Im}\{\boldsymbol{\alpha}\}^T\right]^T \tag{28}$$

The elements of the FIM of a complex circularly Gaussian process $\mathbf{x}(t) \sim \mathcal{N}_C\left(\boldsymbol{\mu}(t), \boldsymbol{R}\right)$ are given by [21], [22],

$$[\boldsymbol{F}]_{l,k} = N \,\mathrm{trace}\left\{\boldsymbol{R}^{-1}\frac{\partial \boldsymbol{R}}{\partial \psi_l}\boldsymbol{R}^{-1}\frac{\partial \boldsymbol{R}}{\partial \psi_k}\right\} + 2\,\mathrm{Re}\left\{\sum_{t=0}^{N-1}\frac{\partial \boldsymbol{\mu}^H(t)}{\partial \psi_l}\boldsymbol{R}^{-1}\frac{\partial \boldsymbol{\mu}(t)}{\partial \psi_k}\right\} \tag{29}$$

Applying (29) to the model (22), we obtain

$$[\boldsymbol{F}]_{l,k} = \frac{NM}{\sigma^4}\frac{\partial \sigma^2}{\partial \psi_l}\frac{\partial \sigma^2}{\partial \psi_k} + \frac{2}{\sigma^2}\,\mathrm{Re}\left\{\sum_{t=0}^{N-1}\frac{\partial\left\{\boldsymbol{\alpha}^H \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial \psi_l}\frac{\partial\left\{\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{\alpha}\right\}}{\partial \psi_k}\right\} \tag{30}$$

Using (30), the following explicit expressions for the blocks of the FIM are derived (see

Appendix A for details),

$$\boldsymbol{F}_{\boldsymbol{\theta\theta}} = \frac{2}{\sigma^2}\mathrm{Re}\left\{\sum_{t=0}^{N-1}\boldsymbol{\Delta}^H\tilde{\boldsymbol{D}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\tilde{\boldsymbol{D}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{\Delta}\right\} \tag{31}$$

$$\boldsymbol{F}_{\boldsymbol{\theta\breve{\alpha}}} = \frac{2}{\sigma^2}\mathrm{Re}\left\{\sum_{t=0}^{N-1}\boldsymbol{\Delta}^H\tilde{\boldsymbol{D}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{Q}\right\} \tag{32}$$

$$\boldsymbol{F}_{\boldsymbol{\theta\nu}_k} = \frac{2}{\sigma^2}\mathrm{Re}\left\{\sum_{t=0}^{N-1}\boldsymbol{\Delta}^H\tilde{\boldsymbol{D}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\left\{\boldsymbol{T}_k(t)\odot\left(\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{\Delta}\right)\right\}\right\} \tag{33}$$

$$\boldsymbol{F}_{\breve{\boldsymbol{\alpha}}\breve{\boldsymbol{\alpha}}} = \frac{2}{\sigma^2}\mathrm{Re}\left\{\sum_{t=0}^{N-1}\boldsymbol{Q}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{Q}\right\} \tag{34}$$

$$\boldsymbol{F}_{\breve{\boldsymbol{\alpha}}\boldsymbol{\nu}_k} = \frac{2}{\sigma^2}\mathrm{Re}\left\{\sum_{t=0}^{N-1}\boldsymbol{Q}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\left\{\boldsymbol{T}_k(t)\odot\left(\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{\Delta}\right)\right\}\right\} \tag{35}$$

$$\boldsymbol{F}_{\boldsymbol{\nu}_k\boldsymbol{\nu}_m} = \frac{2}{\sigma^2}\mathrm{Re}\left\{\sum_{t=0}^{N-1}\left\{\left(\boldsymbol{\Delta}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right)\odot\boldsymbol{T}_k^H(t)\right\}\left\{\boldsymbol{T}_m(t)\odot\left(\tilde{\boldsymbol{A}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{\Delta}\right)\right\}\right\} \tag{36}$$

$$F_{\sigma^2\sigma^2} = \frac{NM}{\sigma^4} \tag{37}$$

$$\boldsymbol{F}_{\boldsymbol{\theta}\sigma^2} = \boldsymbol{F}_{\breve{\boldsymbol{\alpha}}\sigma^2} = \boldsymbol{F}_{\boldsymbol{\nu}_k\sigma^2} = \boldsymbol{0} \tag{38}$$

where

$$\boldsymbol{\Delta} \triangleq \mathrm{diag}\left\{\alpha_1,\cdots,\alpha_L\right\} \tag{39}$$

$$\tilde{\boldsymbol{D}}(\boldsymbol{\theta},\boldsymbol{\omega},t) \triangleq \boldsymbol{D}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{G}(\boldsymbol{\omega},t) \tag{40}$$

$$\boldsymbol{D}(\boldsymbol{\theta},\boldsymbol{\omega},t) \triangleq \left[\frac{\partial\boldsymbol{a}(\theta_1,\boldsymbol{\omega}_1,t)}{\partial\theta_1},\frac{\partial\boldsymbol{a}(\theta_2,\boldsymbol{\omega}_2,t)}{\partial\theta_2},\ldots,\frac{\partial\boldsymbol{a}(\theta_L,\boldsymbol{\omega}_L,t)}{\partial\theta_L}\right] \tag{41}$$

$$\boldsymbol{Q} \triangleq [\boldsymbol{I},j\boldsymbol{I}] \tag{42}$$

$$\boldsymbol{T}_k(t) \triangleq jt^k\left(\frac{t}{k+1}\boldsymbol{E}+\boldsymbol{u}\boldsymbol{c}^T\right) \tag{43}$$

$$\boldsymbol{u} \triangleq \frac{1}{c\,\Delta t}\left[0,d_1,\ldots,d_{M-1}\right]^T \tag{44}$$

$$\boldsymbol{c} \triangleq \left[\sin\theta_1,\sin\theta_2,\ldots,\sin\theta_L\right]^T \tag{45}$$

$\boldsymbol{E}$ is the $M\times L$ matrix containing ones in all positions, and $\boldsymbol{0}$ is the vector of zeros.

With (31)-(38), the FIM has the following structure:

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{\mathcal{F}} & \mathbf{0} \\ \mathbf{0}^T & F_{\sigma^2\sigma^2} \end{bmatrix} \tag{46}$$

where

$$\boldsymbol{\mathcal{F}} = \begin{bmatrix} \boldsymbol{F}_{\theta\theta} & \boldsymbol{F}_{\theta\breve{\alpha}} & \boldsymbol{F}_{\theta\nu_0} & \cdots & \boldsymbol{F}_{\theta\nu_{K-1}} \\ \boldsymbol{F}_{\theta\breve{\alpha}}^T & \boldsymbol{F}_{\breve{\alpha}\breve{\alpha}} & \boldsymbol{F}_{\breve{\alpha}\nu_0} & \cdots & \boldsymbol{F}_{\breve{\alpha}\nu_{K-1}} \\ \boldsymbol{F}_{\theta\nu_0}^T & \boldsymbol{F}_{\breve{\alpha}\nu_0}^T & \boldsymbol{F}_{\nu_0\nu_0} & \cdots & \boldsymbol{F}_{\nu_0\nu_{K-1}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{F}_{\theta\nu_{K-1}}^T & \boldsymbol{F}_{\breve{\alpha}\nu_{K-1}}^T & \boldsymbol{F}_{\nu_0\nu_{K-1}}^T & \cdots & \boldsymbol{F}_{\nu_{K-1}\nu_{K-1}} \end{bmatrix} \tag{47}$$

From (47) and the partitioned matrix inversion formula [22], it follows that

$$\mathbf{CRB} \triangleq \boldsymbol{F}^{-1} = \begin{bmatrix} \boldsymbol{\mathcal{F}}^{-1} & \mathbf{0} \\ \mathbf{0}^T & F_{\sigma^2\sigma^2}^{-1} \end{bmatrix} \tag{48}$$

Equations (31)-(38), (47), and (48) determine an explicit form of the CRB matrix as a function of the unknown source parameters.

V. Chirp Beamformer

In the general case, the implementation of the ML estimator (20)-(21) involves a highly nonlinear optimization over the unknown parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$. Therefore, the associated computational cost may not be always acceptable. In this section, we simplify the ML estimator (20)-(21) by deriving the so-called *chirp beamformer* which requires a simpler 3D search instead of global optimization. Assuming the single source case[3], we have that the matrix $\tilde{\boldsymbol{A}}(\boldsymbol{\theta}, \boldsymbol{\omega}, t)$ simplifies to the vector $\tilde{\boldsymbol{a}}(\boldsymbol{\theta}, \boldsymbol{\omega}, t)$, and the LL function (20) can be written as

$$\begin{aligned} \mathcal{L}_P(\theta_1, \boldsymbol{\omega}_1) &= \left\{ \sum_{t=0}^{N-1} \tilde{\boldsymbol{a}}^H(\theta_1, \boldsymbol{\omega}_1, t) \tilde{\boldsymbol{a}}(\theta_1, \boldsymbol{\omega}_1, t) \right\}^{-1} \left| \sum_{t=0}^{N-1} \boldsymbol{x}^H(t) \tilde{\boldsymbol{a}}(\theta_1, \boldsymbol{\omega}_1, t) \right|^2 \\ &= \frac{1}{NM} \left| \sum_{t=0}^{N-1} \boldsymbol{x}^H(t) \tilde{\boldsymbol{a}}(\theta_1, \boldsymbol{\omega}_1, t) \right|^2 \end{aligned} \tag{49}$$

[3]This assumption will be relaxed later.

where

$$\tilde{\boldsymbol{a}}(\theta_1, \boldsymbol{\omega}_1, t) = g(\boldsymbol{\omega}_1, t)\boldsymbol{a}(\theta_1, \boldsymbol{\omega}_1, t) \tag{50}$$

and the obvious property $\tilde{\boldsymbol{a}}^H\tilde{\boldsymbol{a}} = M$ is used. Assuming a linear FM (chirp) signal, we have $\boldsymbol{\omega}_1 = [\omega_{1,0}, \omega_{1,1}]^T$ and, hence, there are only three parameters $\{\theta_1, \omega_{1,0}, \omega_{1,1}\}$, which correspond to the DOA, initial frequency, and the linear chirp rate, respectively.

Introducing the simplified (subscript-free) notation

$$\theta = \theta_1\,, \qquad \xi = \omega_{1,0}\,, \qquad \zeta = \omega_{1,1} \tag{51}$$

and omitting the constant factor $1/M$, we can rewrite the right-hand side of (49) as the following function:

$$f(\theta, \xi, \zeta) = \frac{1}{N}\left|\sum_{t=0}^{N-1} \boldsymbol{x}^H(t)\tilde{\boldsymbol{a}}(\theta, \xi, \zeta, t)\right|^2 \tag{52}$$

Hereafter, (52) will be referred to as the *chirp beamformer*[4].

The parameters of interest can be obtained from the main maxima of (52) by means of a 3D search over the parameters $\{\theta, \xi, \zeta\}$. It is worth noting that this is a much more feasible problem than the global maximization of the LL function over $L(K+1)$ parameters. The chirp beamformer (52) can be easily applied to the multiple source case under the condition that the sources are "well-separated" in one or more parameters in (51). This property follows from the structure of the function (52), which is linear with respect to the second-order moments of $\boldsymbol{x}$. Therefore, as in the case of the conventional beamformer [23], [25] which is widely used in narrowband array processing, the chirp beamformer (52) can be straightforwardly extended to the multiple source case.

Interestingly, the chirp beamformer has quite a different structure as compared to the

[4]We use this term because of the obvious analogy with the narrowband conventional beamformer [23] which can be easily derived from the conventional deterministic ML estimator under the single-source assumption [24].

conventional beamformer. The latter function is given by [23]

$$\begin{aligned} f_{\mathrm{CB}}(\theta) &= \boldsymbol{a}^H(\theta)\hat{\boldsymbol{R}}\boldsymbol{a}(\theta) \\ &= \frac{1}{N}\sum_{t=0}^{N-1}\left|\boldsymbol{x}^H(t)\boldsymbol{a}(\theta)\right|^2 \end{aligned} \tag{53}$$

where

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{t=0}^{N-1}\boldsymbol{x}(t)\boldsymbol{x}^H(t) \tag{54}$$

is the sample covariance matrix, and the steering vector does not depend on the temporal index t. Comparing (52) and (53), we maintain that the conventional beamformer represents the *sum of the squared absolute values* of vector inner products, whereas the chirp beamformer, on the other hand, is determined by the *squared absolute value of the sum* of inner products. This essential difference between (52) and (53) can be explained by the fact that in the chirp signal case, the signal temporal characteristics are taken into account by means of the parametric time-domain polynomial-phase model. Obviously, this corresponds to the so-called *coherent matched time-domain processing*, whereas in the conventional narrowband case the snapshots $\boldsymbol{x}(t)$ are assumed to be independent and, therefore, the processing in (53) remains *incoherent* in time-domain.

In the single-sensor case, the function (52) becomes independent of θ and yields

$$f(\xi,\zeta) = \frac{1}{N}\left|\sum_{t=0}^{N-1}x(t)e^{-j\xi t}e^{-j\zeta t^2/2}\right|^2 \tag{55}$$

Note that (55) is frequently used for single-channel chirp signal analysis and is commonly referred to as the chirp/quadratic-phase transform [4], [26].

An interesting relationship of the chirp beamformer (52) to traditional estimation techniques can be obtained for the conventional harmonic signal case ($\zeta = 0$). In this case,

$$\tilde{\boldsymbol{a}}(\theta,\xi,\zeta,t) = \tilde{\boldsymbol{a}}(\theta,\xi,t) = e^{j\xi t}\,\boldsymbol{a}(\theta,\xi) \tag{56}$$

where the vector $\boldsymbol{a}(\theta,\xi)$ is the conventional steering vector, which is identical to that in

(53). Hence, the beamforming function (52) can be transformed to

$$\begin{aligned} f(\theta,\xi) &= \left|\boldsymbol{X}^H(\xi)\boldsymbol{a}(\theta,\xi)\right|^2 \\ &= \boldsymbol{a}^H(\theta,\xi)\boldsymbol{X}(\xi)\boldsymbol{X}^H(\xi)\boldsymbol{a}(\theta,\xi) \end{aligned} \tag{57}$$

where

$$\boldsymbol{X}(\xi) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} \boldsymbol{x}(t)\, e^{-j\xi t} \tag{58}$$

is the $M \times 1$ vector of the Fourier-transformed array outputs. The estimator (57) represents a single-snapshot variant of the frequency-domain conventional beamformer [25]

$$f_{\mathrm{CB}}(\theta,\xi) = \boldsymbol{a}^H(\theta,\xi)\hat{\boldsymbol{\mathcal{R}}}(\xi)\boldsymbol{a}(\theta,\xi) \tag{59}$$

where

$$\hat{\boldsymbol{\mathcal{R}}}(\xi) = \frac{1}{P} \sum_{\tau=0}^{P-1} \boldsymbol{X}(\xi,\tau)\boldsymbol{X}^H(\xi,\tau) \tag{60}$$

is the sample spectral density matrix, the time index τ indicates the location of the respective short Fourier transform sliding window[5], and P is the total number of sliding windows (or, in the other words, the number of frequency-domain snapshots).

Similarly to the chirp beamformer (52), a *polynomial-phase beamformer* can be defined that corresponds to a more general PPS. In this case, the number of parameters in (52) will increase, depending on the polynomial-phase model order.

VI. Simulation Results

In all our simulation examples, we assume a uniform linear array (ULA) with omni-directional sensors. The sonar case [27] is considered with the sound propagation speed $c = 1500$ m/s, the interelement spacing $d = 1.5$ m, and the sampling interval $\Delta t = 0.01$ s. The additive white noise is modeled to have identical variances in each array sensor.

[5]This index is not shown in (57) because it is a particular case where the single window, whose length is equal to the whole observation length, is used.

In each simulation example, the experimental DOA and frequency parameter estimation root-mean-square errors (RMSE's) are computed using 50 independent simulation runs and results for both the ML estimator (21) and the chirp beamformer (52) are displayed. The Genetic Algorithm (GA) was used to find the global maximum of the LL function (20). There was no specific initialization of the parameter values. GA is known to be a globally convergent technique and we established the global convergence property through our simulations by a proper choice of the GA parameters, so that there are no outliers caused by the algorithm itself.

It is interesting to compare DOA estimation results for the derived estimators and CRB with the performance which may be expected in the case when *no parametric signal model in time is used*, i.e. when the signal is modeled as

$$\boldsymbol{x}(t) \sim \mathcal{N}_C(\boldsymbol{A}(\boldsymbol{\theta}, \tilde{\boldsymbol{\omega}}(t))\boldsymbol{s}(t), \sigma^2 \boldsymbol{I}) \tag{61}$$

where $\boldsymbol{s}(t)$ are arbitrary unknown deterministic waveforms. Note that if both $\boldsymbol{\theta}$ and $\tilde{\boldsymbol{\omega}}$ are unknown, the problem becomes nonidentifiable and the corresponding CRB does not exist. However, assuming that the instantaneous signal frequencies are exactly known during the whole observation interval[6], we obtain the optimistic deterministic bound corresponding to the aforementioned case when no signal parameterization in time is used. Hereafter, this bound is referred to as the *benchmark deterministic CRB.*

In all subsequent DOA RMSE plots, two CRB's are displayed. The first bound is computed using equations (31)-(38) and referred to as the PPS CRB, whereas the second bound is the benchmark deterministic CRB introduced above.

In the following three examples, we assume a ULA which receives two equi-powered chirp sources impinging on the array from the directions $\theta_1 = 10°$ and $\theta_2 = 15°$ relative to the broadside and having the initial continuous-time frequencies 420 Hz and 401 Hz, respectively, and the continuous-time chirp rates -24 Hz/s and 24 Hz/s, respectively (the

[6]This is obviously an idealistic assumption but it is used to obtain an optimistic result for the CRB.

latter correspond to the discrete-time chirp rates $\omega_{1,1} = -0.01508$ and $\omega_{2,1} = 0.01508$, respectively).

The Wigner-Ville distribution (WVD) computed for $N = 100$ samples of the sampled source waveforms is shown in Fig. 1. Note that this and the subsequent WVD plots display signal waveforms in their baseband representation, which corresponds to downconverted waveforms[7] with the downconversion frequency 400 Hz. In each simulation run, the initial source amplitudes are fixed and equal to $\alpha_1 = e^{-j\pi/8}$ and $\alpha_2 = e^{j\pi/4}$.

In the first example, we assume that $M = 10$ and SNR = 0 dB. In Fig. 2, the RMSE's of the estimated DOA's are shown along with the two aforementioned CRB's versus the SNR. The latter parameter is defined as $\text{SNR} = -10 \log \sigma^2$. Fig. 3 displays the RMSE's of the estimates of the discrete-time initial frequency and chirp rate along with the corresponding PPS CRB's versus the SNR.

From Fig. 2 it follows that in the direction estimation case, the PPS CRB is significantly lower than the benchmark deterministic CRB and the performance of the ML estimator remains very close to the PPS CRB. The RMSE of the chirp beamformer visibly exceeds the PPS CRB because in the multiple signal case this technique is only an approximate solution to the ML estimation problem. At the same time, from Fig. 2 we see that the RMSE of the chirp beamformer is essentially lower than the benchmark deterministic CRB. This important observation demonstrates the role of the matched polynomial-phase processing in time-domain and validates essential improvements in DOA estimates which can be achieved if chirp signals are treated in a correct way. In particular, this comparison clearly shows that the chirp beamformer is a reasonable approximation of the exact ML method. More generally speaking, by exploiting the polynomial-phase temporal signal structure, we are able to improve essentially the DOA estimation accuracy as compared to the techniques which do not exploit such a structure or use a mismatched signal model.

[7]According to a common convention, all time-frequency distributions are displayed in terms of the discrete-time f rather than $\omega = 2\pi f$.

Fig. 3 clearly demonstrates that both the ML estimator and chirp beamformer have excellent performances in estimating the frequency parameters. In particular, their RMSE's remain close to the corresponding PPS CRB's. Note again that ML outperforms the chirp beamformer because in the multiple source case considered, the latter technique is only an approximation of the exact ML estimator.

In the second example, the same parameters as in Example 1 are applied except SNR and N. We take $N = 100$ and the performance is examined versus the SNR. For this example, the DOA estimation RMSE's are displayed in Fig. 4, whereas the RMSE's of the discrete-time frequency parameter estimates are shown in Fig. 5. In both these figures, the corresponding CRB's are shown. Similarly to the previous example, the RMSE's of ML in Figs. 4 and 5 are very close to the respective PPS CRB's. However, the performance of the chirp beamformer can be hardly improved when increasing the SNR. The theoretical reason for this behavior is that the chirp beamformer has a significant bias which does not vanish when increasing the SNR. Since the RMSE includes both the variance and bias-related terms, it becomes constant at high SNR's. Note that such a behavior is quite typical for all conventional (low-resolution) beamforming techniques whose performances are mainly determined by the aperture and data lengths. However, it is worth noting that despite such an "error floor" effect, the chirp beamformer outperforms the deterministic benchmark CRB at low SNR's (Fig. 4).

In the third example, we assume the same parameters as in the examples above, but fixed SNR $= 0$ dB and $N = 100$. The performances of the ML estimator and chirp beamformer are tested versus the number of sensors M. The DOA and frequency parameter estimation RMSE's of both these techniques are plotted in Figs. 6 and 7, respectively, along with the corresponding CRB's. From Fig. 6 we see that at low values of M, the DOA RMSE's of both the ML estimator and chirp beamformer are very close to the PPS CRB which, in turn, is much lower than the benchmark deterministic bound. However, at large sensor numbers the RMSE of chirp beamformer may exceed both the PPS and the benchmark

deterministic CRB's while these two bounds almost merge. Fig. 7 shows that the ML estimates of the frequency parameters keep close to the corresponding PPS CRB's. The chirp beamformer estimates of these parameters are, as a rule, noticeably higher, but still provide a reasonable quality of the frequency estimates.

To summarize these three examples, note that the proposed ML estimator has an excellent performance which achieves the corresponding PPS CRB's, even at low SNR's and small numbers of sensors/samples. The performance of the chirp beamformer is not as high, but still remains better than the benchmark deterministic CRB in several practically important scenarios where the SNR is low and the number of sensors is small. Therefore, chirp beamformer appears to be a reasonable approximation of the exact ML technique.

In the fourth example, we assume a ULA of $M = 10$ sensors and two equi-powered chirp sources with the SNR = 0 dB and the DOA's $\theta_1 = \theta_2 = 30°$. Here, $N = 10$ and the initial continuous-time frequency of each source is equal to 412.73 Hz. The continuous-time chirp rates are -127.32 Hz/s and 127.32 Hz/s, respectively, which correspond to the discrete-time chirp rates $\omega_{1,1} = -0.08$ and $\omega_{2,1} = 0.08$, respectively.

The WVD computed for $N = 100$ samples of the signals corresponding to this example is shown in Fig. 8. From the signal parameter values taken it follows that the two signals are modeled to have identical DOA's and initial frequencies – the only difference between them is in their chirp rates. Fig. 9 displays the 2D slice of the 3D chirp beamforming function evaluated at their true frequency. From this figure, we observe that chirp beamformer is able to resolve closely spaced sources (and even sources having the same DOA's and initial frequencies), based solely on the diversity of their chirp rates. This fact demonstrates that there is an essential advantage in parameterizing PPS's in time, as this enables to make the processing coherent and matched in time-domain, and, therefore, to exploit the signal time-frequency diversity.

An important observation is that although the chirp signals used in our simulations may be subject to aliasing in time-domain (for example, see Fig. 8), there is no ambiguity in

their parameter estimates which are obtained using the ML and chirp beamforming methods. To explain this fact, we stress that both these techniques are *space-time processing* methods and there is no aliasing in the spatial domain. Indeed, in all simulation examples the instantaneous continuous-time signal frequencies do not exceed 500 Hz. The latter frequency corresponds to the case of array sensors spaced half-wavelength apart. Therefore, all instantaneous signal frequencies correspond to interelement spacings which always satisfy the condition $d < \lambda/2$ and, hence, *there is no spatial aliasing.* This explains why there is no ambiguity in the resulting DOA and frequency parameter estimates obtained using the space-time processing techniques (21) and (52).

We end up this section with a remark prompted by our additional simulations whose results are not detailed in the interest of brevity. These simulations have shown that in the large sample size case ($N \gg 1$), the PPS direction estimation CRB is practically independent of the initial signal amplitudes. The same is true for the PPS CRB on the frequency parameters.

VII. Conclusions

The problem of direction finding in sensor arrays in the presence of multiple wideband constant-amplitude polynomial-phase signals was addressed. A new form of the maximum likelihood estimator of signal parameters was presented. This algorithm concentrates the problem at hand with respect to the nuisance initial signal amplitude parameters and is based on the optimization of a nonlinear log-likelihood function.

Furthermore, we derived explicit expressions for the exact Cramér-Rao bound on the accuracy of estimating the signal model parameters. The performance of the maximum likelihood algorithm was illustrated by numerical examples and compared to the obtained Cramér-Rao bound.

Since the proposed exact maximum likelihood estimator may be computationally intensive, a simpler approximate technique was presented, originating from the analysis of

the log-likelihood function in the single chirp signal case. As a result, a new form of spatio-temporal matched filter (referred to as the *chirp beamformer* or, more generally, the *polynomial-phase beamformer*) was derived, which is applicable to a wide class of multiple polynomial-phase signal scenarios. Our chirp beamforming approach requires solving a 3D optimization problem and, therefore, enjoys essentially simpler implementation than that associated with the exact maximum likelihood technique. Numerical examples were presented showing the DOA estimation performance of the chirp beamformer against the Cramér-Rao bound for different values of SNR, as well as different numbers of snapshots and array sensors. Although the exact ML estimator outperforms the chirp beamformer, the latter technique appears to be a reasonable approximation of ML.

Appendix A: Derivation of the FIM

Let us obtain closed-form expressions for all particular subblocks of the FIM.

Direct computations yield

$$\frac{\partial\left\{\boldsymbol{\alpha}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial\mathrm{Re}\{\boldsymbol{\alpha}\}} = \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t) \tag{62}$$

$$\frac{\partial\left\{\boldsymbol{\alpha}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial\mathrm{Im}\{\boldsymbol{\alpha}\}} = -j\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t) \tag{63}$$

Introducing the $L \times 2L$ matrix

$$\boldsymbol{Q} \triangleq [\boldsymbol{I}, j\boldsymbol{I}] \tag{64}$$

we can rewrite (62) and (63) in a compact form as

$$\frac{\partial\left\{\boldsymbol{\alpha}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial\breve{\boldsymbol{\alpha}}} = \boldsymbol{Q}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t) \tag{65}$$

Straightforward computations also yield

$$\begin{aligned}\frac{\partial\left\{\boldsymbol{\alpha}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial\theta_l} &= \boldsymbol{\alpha}^H\boldsymbol{G}^H(\boldsymbol{\omega},t)\,\boldsymbol{e}_l\boldsymbol{e}_l^T\boldsymbol{D}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\\ &= \boldsymbol{\alpha}^H\boldsymbol{e}_l\boldsymbol{e}_l^T\boldsymbol{G}^H(\boldsymbol{\omega},t)\,\boldsymbol{D}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\\ &= \boldsymbol{\alpha}^H\boldsymbol{e}_l\boldsymbol{e}_l^T\tilde{\boldsymbol{D}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\end{aligned} \tag{66}$$

where $\boldsymbol{e}_l$ is the lth column of the identity matrix (i.e. vector containing one in the lth position and zeros elsewhere), the diagonal structure of $\boldsymbol{G}(\boldsymbol{\omega},t)$ is used, and the following notations are introduced:

$$\tilde{\boldsymbol{D}}(\boldsymbol{\theta},\boldsymbol{\omega},t) \triangleq \boldsymbol{D}(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{G}(\boldsymbol{\omega},t) \tag{67}$$

$$\boldsymbol{D}(\boldsymbol{\theta},\boldsymbol{\omega},t) \triangleq \left[\frac{\partial\boldsymbol{a}(\theta_1,\boldsymbol{\omega}_1,t)}{\partial\theta_1}, \frac{\partial\boldsymbol{a}(\theta_2,\boldsymbol{\omega}_2,t)}{\partial\theta_2}, \ldots, \frac{\partial\boldsymbol{a}(\theta_L,\boldsymbol{\omega}_L,t)}{\partial\theta_L}\right] \tag{68}$$

From (66), we obtain that

$$\begin{aligned}\frac{\partial\left\{\boldsymbol{\alpha}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial\boldsymbol{\theta}} &= \left[\tilde{\boldsymbol{D}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\,\boldsymbol{e}_1\boldsymbol{e}_1^T\boldsymbol{\alpha}\,,\,\tilde{\boldsymbol{D}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\,\boldsymbol{e}_2\boldsymbol{e}_2^T\boldsymbol{\alpha}\,,\,\ldots\,,\,\tilde{\boldsymbol{D}}(\boldsymbol{\theta},\boldsymbol{\omega},t)\,\boldsymbol{e}_L\boldsymbol{e}_L^T\boldsymbol{\alpha}\right]^H\\ &= \boldsymbol{\Delta}^H\tilde{\boldsymbol{D}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\end{aligned} \tag{69}$$

where

$$\boldsymbol{\Delta} \triangleq \text{diag}\left\{\alpha_1, \cdots, \alpha_L\right\} \tag{70}$$

Let us introduce the following notation

$$\begin{aligned}
h_k\left(\boldsymbol{\omega}_l, t\right) &\triangleq \frac{\partial g\left(\boldsymbol{\omega}_l, t\right)}{\partial \omega_{l,k}} \\
&= j\,\frac{t^{k+1}}{k+1}\exp\left\{j\sum_{s=0}^{K-1}\omega_{l,s}\frac{t^{s+1}}{s+1}\right\} && (71)\\
\boldsymbol{g}(\boldsymbol{\omega}, t) &\triangleq \left[g(\boldsymbol{\omega}_1, t), g(\boldsymbol{\omega}_2, t), \ldots, g(\boldsymbol{\omega}_L, t)\right]^T && (72)\\
\boldsymbol{H}_k\left(\boldsymbol{\omega}, t\right) &\triangleq \frac{\partial \boldsymbol{g}^T\left(\boldsymbol{\omega}, t\right)}{\partial \boldsymbol{\nu}_k} \\
&= \text{diag}\left\{h_k(\boldsymbol{\omega}_1, t), \ldots, h_k(\boldsymbol{\omega}_L, t)\right\} \\
&= j\,\frac{t^{k+1}}{k+1}\,\boldsymbol{G}(\boldsymbol{\omega}, t) && (73)\\
\boldsymbol{b}_{m,k}(\theta_l, \boldsymbol{\omega}_l, t) &\triangleq \frac{\partial \boldsymbol{a}(\theta_l, \boldsymbol{\omega}_l, t)}{\partial \omega_{m,k}} && (74)\\
&= j\,t^k \sin\theta_l\,\boldsymbol{U}\,\boldsymbol{a}(\theta_l, \boldsymbol{\omega}_l, t)\,\delta_{l,m} && (75)\\
\boldsymbol{U} &\triangleq \frac{1}{c\,\Delta t}\,\text{diag}\left\{0, d_1, \ldots, d_{M-1}\right\} && (76)\\
\boldsymbol{C} &\triangleq \text{diag}\left\{\sin\theta_1, \sin\theta_2, \ldots, \sin\theta_L\right\} && (77)\\
\boldsymbol{B}_k(\boldsymbol{\theta}, \boldsymbol{\omega}, t) &\triangleq \left[\frac{\partial \boldsymbol{a}(\theta_1, \boldsymbol{\omega}_1, t)}{\partial \omega_{1,k}}, \frac{\partial \boldsymbol{a}(\theta_2, \boldsymbol{\omega}_2, t)}{\partial \omega_{2,k}}, \ldots, \frac{\partial \boldsymbol{a}(\theta_L, \boldsymbol{\omega}_L, t)}{\partial \omega_{L,k}}\right] && (78)\\
&= j\,t^k\,\boldsymbol{U}\boldsymbol{A}(\boldsymbol{\theta}, \boldsymbol{\omega}, t)\boldsymbol{C} && (79)\\
\boldsymbol{u} &\triangleq \frac{1}{c\,\Delta t}\left[0, d_1, \ldots, d_{M-1}\right]^T && (80)\\
\boldsymbol{c} &\triangleq \left[\sin\theta_1, \sin\theta_2, \ldots, \sin\theta_L\right]^T && (81)\\
\boldsymbol{T}_k(t) &\triangleq jt^k\left(\frac{t}{k+1}\boldsymbol{E} + \boldsymbol{u}\boldsymbol{c}^T\right) && (82)
\end{aligned}$$

where $\boldsymbol{E}$ is the matrix containing ones in all positions and $\delta_{l,m}$ is the Kronecker delta.

Using (71)-(79), we obtain that

$$\begin{aligned}\frac{\partial\left\{\boldsymbol{\alpha}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial\omega_{l,k}} &= \frac{\partial\left\{\boldsymbol{g}^H(\boldsymbol{\omega},t)\boldsymbol{\Delta}^H\boldsymbol{A}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial\omega_{l,k}} \\ &= \frac{\partial\left\{\boldsymbol{g}^H(\boldsymbol{\omega},t)\right\}}{\partial\omega_{l,k}}\boldsymbol{\Delta}^H\boldsymbol{A}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)+\boldsymbol{g}^H(\boldsymbol{\omega},t)\boldsymbol{\Delta}^H\frac{\partial\left\{\boldsymbol{A}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial\omega_{l,k}} \\ &= \boldsymbol{e}_l^T\boldsymbol{H}_k^H(\boldsymbol{\omega},t)\boldsymbol{\Delta}^H\boldsymbol{A}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)+\boldsymbol{g}^H(\boldsymbol{\omega},t)\boldsymbol{\Delta}^H\boldsymbol{e}_l\boldsymbol{e}_l^T\boldsymbol{B}_k^H(\boldsymbol{\theta},\boldsymbol{\omega},t) \quad (83)\end{aligned}$$

From (83), (73), and (78), it follows that

$$\begin{aligned}\frac{\partial\left\{\boldsymbol{\alpha}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial\boldsymbol{\nu}_k} &= \boldsymbol{H}_k^H(\boldsymbol{\omega},t)\boldsymbol{\Delta}^H\boldsymbol{A}^H(\boldsymbol{\theta},\boldsymbol{\omega},t) \\ &\quad+\left[\boldsymbol{B}_k(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{e}_1\boldsymbol{e}_1^T\boldsymbol{\Delta}\boldsymbol{g}(\boldsymbol{\omega},t),\ldots,\boldsymbol{B}_k(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{e}_L\boldsymbol{e}_L^T\boldsymbol{\Delta}\boldsymbol{g}(\boldsymbol{\omega},t)\right]^H \\ &= \boldsymbol{H}_k^H(\boldsymbol{\omega},t)\boldsymbol{\Delta}^H\boldsymbol{A}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)+\boldsymbol{G}^H(\boldsymbol{\omega},t)\boldsymbol{\Delta}^H\boldsymbol{B}_k^H(\boldsymbol{\theta},\boldsymbol{\omega},t) \\ &= -j\frac{t^{k+1}}{k+1}\boldsymbol{G}^H(\boldsymbol{\omega},t)\boldsymbol{\Delta}^H\boldsymbol{A}^H(\boldsymbol{\theta},\boldsymbol{\omega},t) \\ &\quad -j\,t^k\boldsymbol{G}^H(\boldsymbol{\omega},t)\boldsymbol{\Delta}^H\boldsymbol{C}\boldsymbol{A}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{U} \quad (84)\end{aligned}$$

Using (80), (81), and exploiting the diagonal structure of the matrices $\boldsymbol{\Delta}$, $\boldsymbol{C}$, and $\boldsymbol{U}$, we can write

$$\boldsymbol{G}^H(\boldsymbol{\omega},t)\boldsymbol{\Delta}^H\boldsymbol{C}\boldsymbol{A}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\boldsymbol{U}=\left(\boldsymbol{\Delta}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right)\odot\left(\boldsymbol{c}\boldsymbol{u}^T\right) \quad (85)$$

where $\odot$ is the Schur-Hadamard (elementwise) matrix product. Inserting (85) into (84) and using (82) gives

$$\frac{\partial\left\{\boldsymbol{\alpha}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right\}}{\partial\boldsymbol{\nu}_k} = \left(\boldsymbol{\Delta}^H\tilde{\boldsymbol{A}}^H(\boldsymbol{\theta},\boldsymbol{\omega},t)\right)\odot\boldsymbol{T}_k^H(t) \quad (86)$$

Using (30), (65), (69), and (86), we obtain the explicit expressions (31)-(38) for the FIM subblocks.

References

[1] Advanced Radar Techniques and Systems, G. Galati, Ed., IEE Press, 1993.

[2] J. Chatillon, M.E. Bouhier, and M.E. Zakharia, "Synthetic aperture sonar for seabed imaging: relative merits of narrowband and wideband approaches," *IEEE Journal Oceanic Engineering*, vol. 17, pp. 95-105, Jan. 1992.

[3] S. Barbarossa, A. Scaglione, and G.B. Giannakis, "Product high-order ambiguity function for multicomponent polynomial-phase signal modeling", *IEEE Trans. Signal Processing*, vol. 46, pp. 691-707, March 1998.

[4] X.-G. Xia, "Discrete chirp-Fourier transform and its application to chirp rate estimation," *IEEE Trans. Signal Processing*, vol. 48, pp. 3122-3133, Nov. 2000.

[5] P.M. Djuric and S.M. Kay, "Parameter estimation of chirp signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 2118-2126, Dec. 1990.

[6] S. Peleg and B. Porat, "Estimation and classification of polynomial-phase signals," *IEEE Trans. Inform. Theory*, vol. 37, pp. 422-430, March 1991.

[7] S. Shamsunder, G.B. Giannakis, and B. Friedlander, "Estimating random-amplitude polynomial phase signals: A cyclostationary approach," *IEEE Trans. Signal Processing*, vol. 43, pp. 492-505, Feb. 1995.

[8] O. Besson, M. Ghogho, and A. Swami, "Parameter estimation for random amplitude chirp signals," *IEEE Trans. Signal Processing*, vol. 47, pp. 3208-3219, Dec. 1999:

[9] A. Zeira and B. Friedlander, "Joint direction finding, signal, and channel response estimation for a polynomial phase signal in a multipath channel," in *Proc. 30th Asilomar Conf. Sign., Syst., Comp.*, Pacific Grove, CA, pp. 733-737, Nov. 1996.

[10] M.G. Amin, "Spatial time-frequency distributions for direction finding and blind source separation," *Proc. SPIE*, vol. 3723, pp. 62-70, 1999.

[11] A. Belouchrani and M.G. Amin, "Time-frequency MUSIC," *IEEE Signal Processing Letters*, vol. 6, pp. 109-110, 1999.

[12] B. Völcker and B. Ottersten, "Linear chirp parameter estimation from multi channel data," in *Proc. 32nd Asilomar Conf. Sign., Syst., Comp.*, Pacific Grove, CA, Nov. 1998.

[13] A.B. Gershman and M.G. Amin, "Coherent wideband DOA estimation of multiple FM signals using spatial time-frequency distributions," in *Proc. ICASSP'00*, Istanbul, Turkey, pp. 3065-3068, June 2000.

[14] A.B. Gershman and M.G. Amin, "Wideband direction of arrival estimation of multiple chirp signals using spatial time-frequency distributions," *IEEE Signal Processing Letters*, vol. 7, pp. 152-155, June 2000.

[15] G. Wang and X.-G. Xia, "Iterative algorithm for direction of arrival estimation with wideband chirp signals," *IEE Proc. Radar, Sonar, Navig.*, vol. 147, pp. 233-238, Oct. 2000.

[16] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 823-831, Aug. 1985.

[17] B. Friedlander and A.J. Weiss, "Direction finding for wide-band signals using an interpolated array," *IEEE Trans. Signal Processing*, vol. 41, pp. 1618-1634, Apr. 1993.

[18] J.F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields," *Signal Processing*, vol. 10, pp. 329-337, 1986.

[19] M. Wax, "Detection and estimation of superimposed signals," Ph.D. Dissertation, Stanford Univ., Stanford, CA, March 1985.

[20] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and Cramér-Rao bound," *IEEE Trans. Acoust., Speech., Signal Processing*, vol. 37, pp. 720-741, May 1989.

[21] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1783-1795, Oct. 1990.

[22] P. Stoica and R.L. Moses, *Introduction to Spectral Analysis*, Prentice-Hall, Upper Saddle River, NJ, 1997.

[23] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Processing Magazine*, vol. 13, pp. 67-94, July 1996.

[24] V. Katkovnik and A.B. Gershman, "A local polynomial approximation based beamforming for source localization and tracking in nonstationary environments," *IEEE Signal Processing Letters*, vol. 7, pp. 3-5, Jan. 2000.

[25] J.F. Böhme, "Array processing", in *Advances in Spectrum Estimation*, S. Haykin, Ed., Prentice-Hall, Englewood Cliffs, NJ, 1991, vol. II, pp. 1-63.

[26] M.Z. Ikram, K. Abed-Meraim, and Y. Hua, "Fast discrete quadratic phase transform for estimating the parameters of chirp signals," in *Proc. 30th Asilomar Conf. Sign., Syst., Comp.*, Pacific Grove, CA, pp. 798 -802, Nov. 1996.

[27] A.B. Gershman, V.I. Turchin, and V.A. Zverev, "Experimental results of localization of moving underwater signal by adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 43, pp. 2249-2257, Oct. 1995.

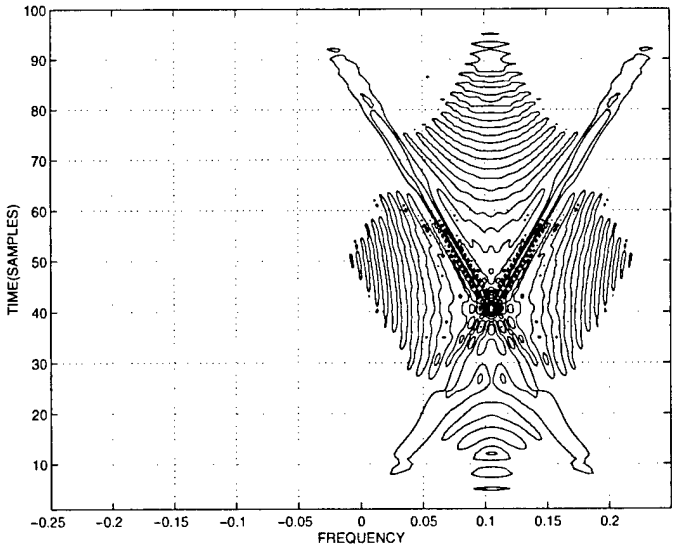


Fig. 1. WVD of the source waveforms. Examples 1-3.

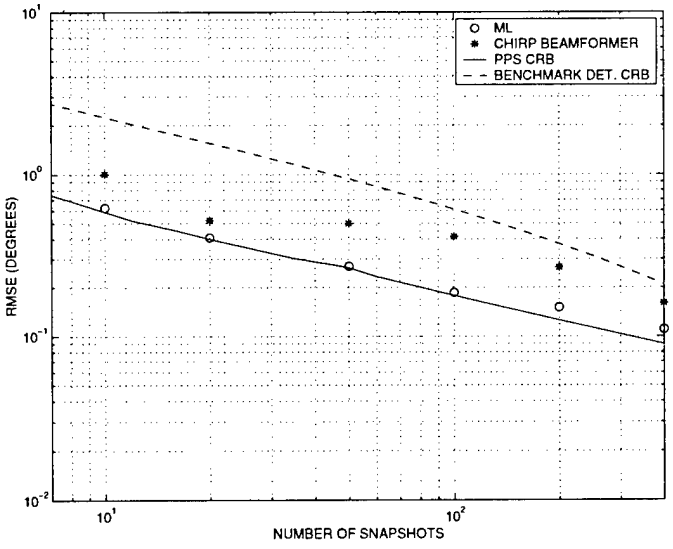


Fig. 2. Comparison of the DOA estimation RMSE's of the ML estimator and chirp beamformer with the PPS and benchmark deterministic CRB's versus the number of snapshots. First example.

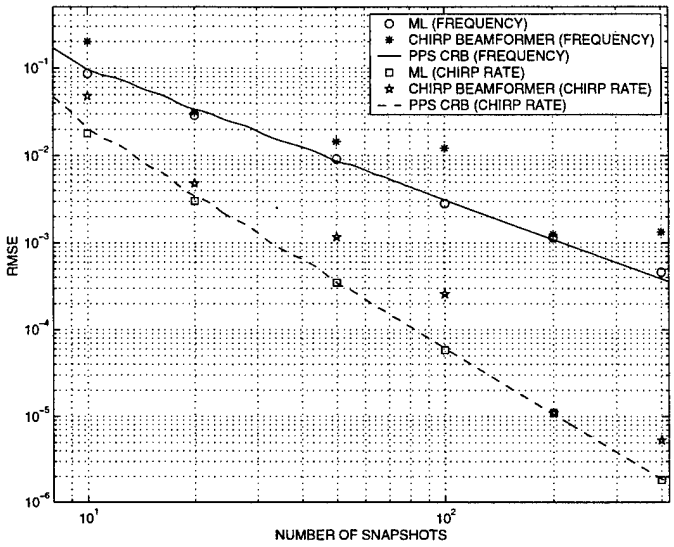


Fig. 3. Comparison of the frequency parameter estimation RMSE's of the ML estimator and chirp beamformer with the PPS CRB's versus the number of snapshots. First example.

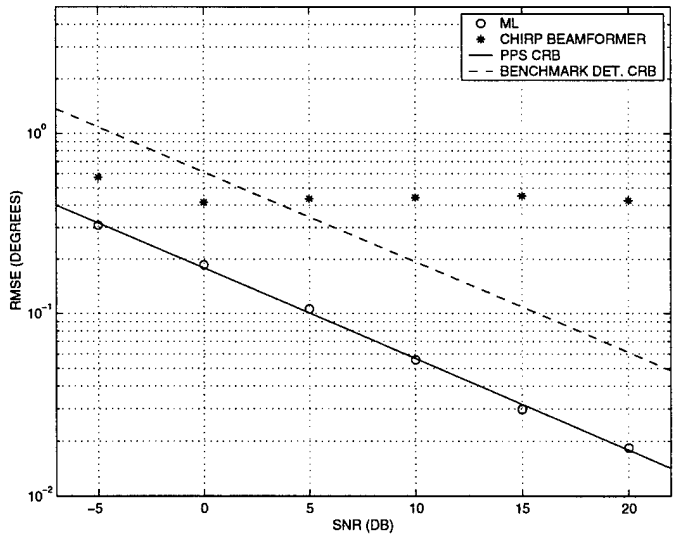


Fig. 4. Comparison of the DOA estimation RMSE's of the ML estimator and chirp beamformer with the PPS and benchmark deterministic CRB's versus the SNR. Second example.

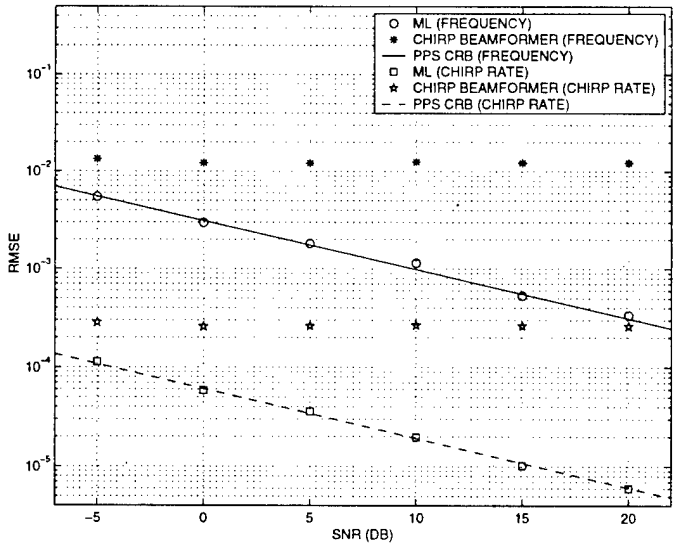


Fig. 5. Comparison of the frequency parameter estimation RMSE's of the ML estimator and chirp beamformer with the PPS CRB's versus the SNR. Second example.

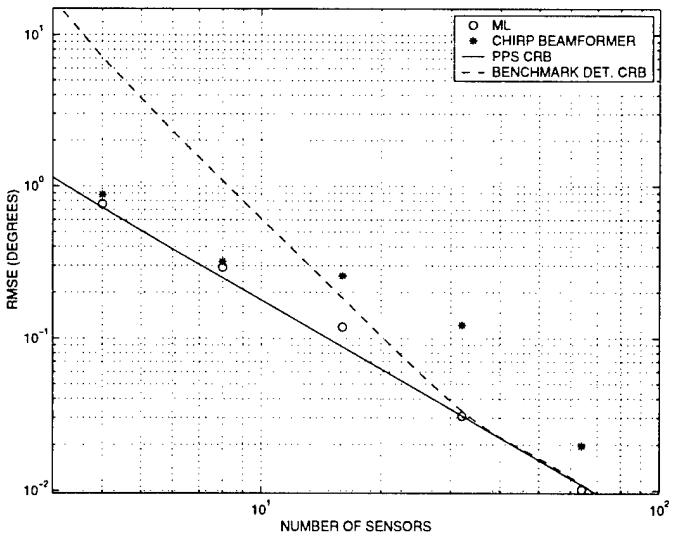


Fig. 6. Comparison of the DOA estimation RMSE's of the ML estimator and chirp beamformer with the PPS and benchmark deterministic CRB's versus the number of sensors. Third example.

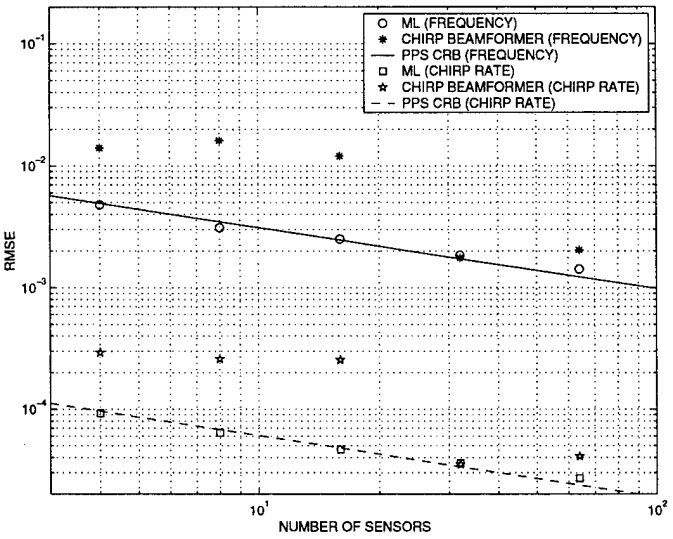


Fig. 7. Comparison of the frequency parameter estimation RMSE's of the ML estimator and chirp beamformer with the PPS CRB's versus the number of sensors. Third example.

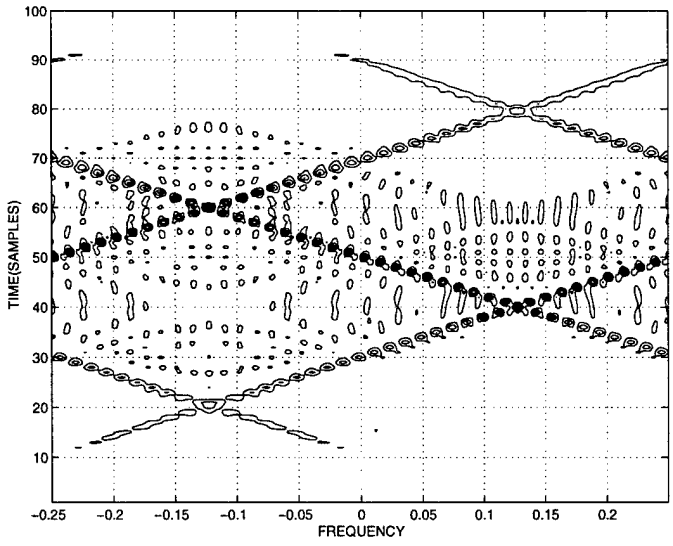


Fig. 8. WVD of the source waveforms. Fourth example.

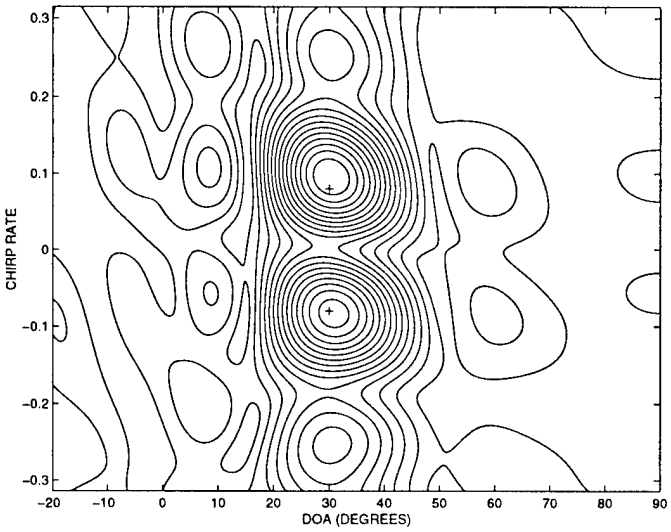


Fig. 9. 2D slice of the chirp beamformer. The true source locations are indicated + signs. Fourth example.

Performance Analysis of Subband Arrays

Yimin Zhang[†], Kehu Yang[‡], Moeness G. Amin[†], and Yoshio Karasawa[§]
[†] Department of Electrical and Computer Engineering,
Villanova University, Villanova, PA 19085
[‡] ATR Adaptive Communications Research Laboratories,
Kyoto 619-0288, Japan
[§] Department of Electronic Engineering,
University of Electro-Communications, Tokyo 182-8585, Japan

Abstract

Several subband array methods have been proposed as useful means to perform joint spatio-temporal equalization in digital mobile communications. These methods can be applied to mitigate problems caused by the inter-symbol interference (ISI) and co-channel interference (CCI). The subband array methods proposed so far can be classified into two major schemes: (1) a centralized feedback scheme and (2) a localized feedback scheme. In this paper, we propose subband arrays with partial feedback scheme, which generalize the above two feedback schemes.

The main contribution of this paper is to derive the steady-state mean square error (MSE) performance of subband arrays implementing these three different feedback schemes. Unlike the centralized feedback scheme which can be designed to provide the optimum equalization performance, the subband arrays with localized and partial feedback schemes are in general suboptimal. The performance of these two suboptimal feedback schemes depends on the channel characteristics, the filter banks employed, and the number of subbands.

Keywords

Subband array, space-time adaptive processing, adaptive array, multirate signal processing, mobile communications.

The work of Y. Zhang is continued from his previous work when he was with the ATR Adaptive Communications Research Laboratories. The work of M. G. Amin is supported by the Office of Naval Research under Grant N00014-98-1-0176.

I. Introduction

Mobile communication systems are developing toward higher-speed digital wireless networks. Their applications are rapidly expanding from voice transmission to a wide class of multimedia information. In the new wireless networks, the communication channels are often frequency-selective, which makes the inter-symbol interference (ISI) to be highly pronounced. Another important problem in mobile communication is the co-channel interference (CCI), which is the result of frequency reuse in cellular systems.

Adaptive arrays implementing spatial or spatio-temporal equalizations prove useful in suppressing both ISI and CCI, leading to improved communication quality and increased communication capacity [1]–[4]. Specifically, space-time adaptive processing (STAP) techniques are power tools to achieve spatio-temporal equalizations. The high complexity and slow convergence, however, are key issues in practical implementation of STAP systems.

Recently, subband adaptive array methods have been proposed as alternative tools for spatio-temporal equalization. The authors have proposed in [5] – [8] to use subband arrays to realize joint spatio-temporal equalizations. This concept has also been extended to subband STAP schemes [9], [10]. Compared with conventional STAP systems, subband adaptive arrays offer amenability to parallel implementations [8], rapid convergence [11], [12], and a reduction of processing complexity [13], [14]. Subband processing is cast in [15] as an elegant and computationally efficient solution to the needs for increased bandwidth in array processing applications.

The subband array methods proposed so far can be classified, in terms of the definition of error signals used to control the weight updation, into two major classes: (1) a centralized feedback scheme and (2) a localized feedback scheme. A subband array with the localized feedback scheme allows parallel subband processing with greatly reduced computations at each subband, accompanied with improved convergence. These features are vert attractive in STAP implementations, as the system complexity increases sharply when either or all

of the data rate, delay profile, and the number of array sensors increase.

We propose in this paper the partial feedback scheme, which generalizes the above two feedback schemes. The proposed partial feedback scheme permits more flexibility in trading-off the system complexity, converegence, and the steady-state mean square error (MSE) performance.

Our main contribution in this paper is analysis of the MSE performance of subband arrays with the three different feedback schemes. For simplicity of analysis and comparison, it is assumed that the reference signal is available. For the centralized feedback schemes, reference [16] has shown that frequency domain array processing provides the same steady-state MSE performance as that offered by the STAP system, using tapped delay-lines (TDL). Reference [17] provides important comparison results between the centralized and localized feedback schemes. However, such comparison was limited to the simulation results, and analytical support was not presented.

In this paper, we consider the analytical results of MSE performance of subband arrays with the three different feedback schemes. To the best of our knowledge, such results for the localized and partial feedback schemes have not yet been produced. It is shown in the following discussion that, unlike the centralized feedback subband array, which gives the optimum spatio-temporal equalization performance, the MSE performance provided by the localized and partial feedback subband arrays are generally suboptimal. The performance of these two suboptimal feedback schemes depends on the channel characteristics, the filter banks employed, and the number of subbands.

This paper is organized as follows. In Section 2, we introduce the signal model, and the steady-state MSE performance of the STAP systems is described. In Section 3, the subband decomposition is introduced, and the steady-state MSE performance of the centralized feedback subband array is derived and shown to be equivalent to the optimum STAP results. Section 4 analyzes the steady-state MSE performance of localized feedback subband arrays. In Section 5, the partial feedback scheme is proposed and its steady-state

MSE performance is analyzed. Section 6 provides simulation examples for the covariance matrices of the original and the subband signals. The MSE results are compared for different feedback schemes.

II. Signal Model

We consider a base station that uses an antenna array of N sensors with P users, where $P < N$. The signal of interest is denoted by $s_1(l), l \in (-\infty, \infty)$, whereas the signals from the other users are denoted by $s_p(l)$, $p = 2, ..., P$. Accordingly, the received signal vector $\vec{x}(l)$ at the array, expressed in discrete form, is given by

$$\vec{x}(l) = \sum_{p=1}^{P} \sum_{m=-\infty}^{\infty} s_p(m)\vec{h}_p(l-m) + \vec{b}(l) \tag{1}$$

where

$s_p(l)$: information symbol of the pth user,

$\vec{h}_p(l)$: channel response vector of the pth user,

$\vec{b}(l)$: additive noise vector.

In this paper, we restrict the discussion to T-spaced equalization (i.e., sampled at the symbol rate) for simplicity. We make the following assumptions.

A1) The user signals $s_p(l), p = 1, 2, ..., P$, are wide-sense stationary and independent and identically distributed (i. i. d.) with $E[s_p(l)s_p^*(l)] = 1$, where the superscript * denotes complex conjugate.

A2) All channels $\vec{h}_p(l), p = 1, 2, ..., P$, are linear time-invariant and of a finite duration within $[0, D_p]$. That is, $\vec{h}_p(l) = 0, p = 1, 2, ..., P$, for $l > D_p$ and $l < 0$.

A3) The noise vector $\vec{b}(l)$ is zero-mean, temporally and spatially white with

$$E[\vec{b}(l)\vec{b}^T(l)] = \mathbf{0}, \quad \text{and} \quad E[\vec{b}(l)\vec{b}^H(l)] = \sigma \mathbf{I}_N,$$

where the superscripts ^{T} and ^{H} denote transpose and conjugate transpose, respectively, σ is the noise power, and $\mathbf{I}_N$ is the $N \times N$ identity matrix.

Considering M successive snapshots, we have

$$\mathbf{x}(l) = \sum_{p=1}^{P} \mathbf{H}_p \mathbf{s}_p(l) + \mathbf{b}(l) \tag{2}$$

where

$$\mathbf{x}(l) = [\,\vec{x}^T(l)\ \vec{x}^T(l-1)\ \cdots\ \vec{x}^T(l-M+1)\,]^T \tag{3}$$

$$\mathbf{H}_p = \begin{bmatrix} \vec{h}_p(0) & \cdots & \vec{h}_p(D_p) & 0 & \cdots & \cdots & 0 \\ 0 & \vec{h}_p(0) & \cdots & \vec{h}_p(D_p) & 0 & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & \cdots & 0 & \vec{h}_p(0) & \cdots & \vec{h}_p(D_p) \end{bmatrix} \tag{4}$$

$$\mathbf{s}_p(l) = [\,s_p(l)\ s_p(l-1)\ \cdots\ s_p(l-M-D_p+1)\,]^T \tag{5}$$

and

$$\mathbf{b}(l) = [\,\vec{b}^T(l) \quad \vec{b}^T(l-1) \quad \cdots \quad \vec{b}^T(l-M+1)\,]^T . \tag{6}$$

Denote $\vec{w}(m)$ as the weight vector of the STAP system corresponding to $\vec{x}(l-m)$, and define $\mathbf{w}(l) = \left[\vec{w}^T(l), \cdots, \vec{w}^T(l-M+1)\right]^T$. Then, the output of the STAP becomes

$$y(l) = \mathbf{w}^T(l)\mathbf{x}(l) = \sum_{m=0}^{M-1} \vec{w}^T(m)\vec{x}(l-m). \tag{7}$$

Using the minimum mean square error (MMSE) criterion,

$$\min_{\mathbf{w}} E\,|y(l) - s_1(l-v)|^2 = \min_{\mathbf{w}} E\left|\mathbf{w}^T\mathbf{x}(l) - s_1(l-v)\right|^2 \tag{8}$$

where $0 \le v \le M + D_1 - 1$ is an appropriate time delay which minimizes the MSE [10], then the optimum weight vector is given by the Weiner-Hopf solution

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{r} \tag{9}$$

where

$$\mathbf{R} = E[\mathbf{x}^*(l)\mathbf{x}^T(l)] \tag{10}$$

is the correlation matrix of $\mathbf{x}(l)$, and

$$\mathbf{r} = E[\mathbf{x}^*(l)s_1(l-v)] \tag{11}$$

is the cross-correlation vector between $\mathbf{x}(l)$ and the training signal, which is assumed to be an ideal replica of $s_1(l)$. The superscript * denotes complex conjugate. Substituting (2) to (11) yields

$$\mathbf{r} = E\left[\left(\sum_{p=1}^{P}\mathbf{H}_p\mathbf{s}_p(l)+\mathbf{b}(l)\right)^* s_1(l-v)\right] = E\left[\mathbf{H}_1^*\mathbf{s}_1^*(l)s_1(l-v)\right] = \mathbf{H}_1^*\mathbf{e}_{v+1}, \tag{12}$$

where $\mathbf{e}_{v+1} = [0 \cdots 0\ 1\ 0 \cdots 0]^T$ is a vector whose elements are zero except that at the $v+1$ element being 1. It is obvious that $\mathbf{r}$ is the $(v+1)$-th column of $\mathbf{H}_1^*$.

Since $\mathbf{R}$ is Hermitian, then the MMSE is given by

$$\begin{aligned}
\text{MMSE} &= E\left|\mathbf{w}_{opt}^T\mathbf{x}(l) - s_1(l)\right|^2 \\
&= E\left|\mathbf{r}^T(\mathbf{R}^{-1})^T\mathbf{x}(l) - s_1(l)\right|^2 \\
&= \mathbf{r}^T(\mathbf{R}^{-1})^T E[\mathbf{x}(l)\mathbf{x}^H(l)](\mathbf{R}^{-1})^*\mathbf{r}^* \\
&\quad -\mathbf{r}^T(\mathbf{R}^{-1})^T E[\mathbf{x}(l)s_1^*(l)] \\
&\quad +E[s_1(l)s_1^*(l)] \\
&\quad -\mathbf{r}^H\mathbf{R}^{-1}E[\mathbf{x}^*(l)s_1(l)] \\
&= 1 - \mathbf{r}^H\mathbf{R}^{-1}\mathbf{r}.
\end{aligned} \tag{13}$$

III. Subband Arrays

A. Subband Decomposition

Subband decomposition is performed by exploiting a set of analysis and synthesis filters. Discrete Fourier transform (DFT) and modified-QMF filter banks are examples of perfect reconstructed (PR) and near-perfect reconstruction (NPR) filter banks, respectively [8]. Decimation can be applied between the analysis filters and the synthesis filters to reduce the processing data rate. The decimation rate should not exceed the number of subbands.

Such decimation, however, often reduces the steady state system performance due to aliasing. We maintain that, the PR and NPR properties can be easily destroyed if adaptive techniques are employed between the analysis filters and the synthesis filters because of the changes in the aliasing characteristics. In this paper, no decimation is performed for subband signal components. In this case, the synthesis filters are either not necessary, or can be integrated at the analysis filters.

Let the subband decomposition devide the data sequence at the output of ith virtual channel, $\tilde{x}_i(l)$, into Q subband sequences, $x_i^{(1)}(l), \cdots, x_i^{(Q)}(l)$, where the superscript (m) denotes the signal component at the mth subband. We define

$$\mathbf{x}_T(l) = \left[\left(\vec{x}_T^{(1)}(l)\right)^T, \cdots, \left(\vec{x}_T^{(Q)}(l)\right)^T\right]^T$$

as the signal vector for the subband arrays with

$$\vec{x}_T^{(m)}(l) = \left[x_1^{(m)}(l), x_2^{(m)}(l) \cdots, x_N^{(m)}(l)\right]^T.$$

As a general expression, we can relate $\mathbf{x}_T(l)$ and $\mathbf{x}(l)$ by a $QN \times MN$ transform matrix as

$$\mathbf{x}_T(l) = \mathbf{T}\mathbf{x}(l). \tag{14}$$

We only consider the specific cases where $\mathbf{T}$ is square (i.e., $Q = M$) and unitary (i.e., $\mathbf{T}\mathbf{T}^H = \mathbf{T}^H\mathbf{T} = \mathbf{I}_{MN}$). That is, the number of subbands is set equal to the number of the snapshots at each array sensor. This kind of subband processing is also known as real-time transform-domain processing [18].

A good example of such transform is the DFT filter bank, where the transform matrix $\mathbf{T}$ can be expressed in the form

$$\mathbf{T} = \mathbf{P}^T(\mathbf{I}_N \otimes \mathbf{T}_o)\mathbf{P} \tag{15}$$

where $\otimes$ denotes Kronecker product, and

$$\mathbf{T}_o = \frac{1}{\sqrt{M}} \begin{bmatrix} W_M^0 & W_M^0 & W_M^0 & \cdots & W_M^0 \\ W_M^0 & W_M^1 & W_M^2 & \cdots & W_M^{M-1} \\ \vdots & & & & \vdots \\ W_M^0 & W_M^{M-1} & W_M^{2(M-1)} & \cdots & W_M^{(M-1)^2} \end{bmatrix} \tag{16}$$

with $W_M = \exp\left(\frac{-2\pi j}{M}\right)$. In (15), $\mathbf{P}$ is a permutation matrix to change the order of the elements of vector $\mathbf{x}(l)$ such that the M samples at each array sensor align together.

The DFT filter bank satisfies the PR condition [19] because the only non-zero sum of the column vectors (i.e., the coefficients of the analysis filters for different subbands) of $\mathbf{T}_o$ appears at the first column.

B. Subband Array with Centralized Feedback

In this part, we consider the subband array with centralized feedback scheme, as illustrated in Fig. 1. Weighting $\mathbf{x}_T(l)$ by the weight vector $\mathbf{w}_T = \left[(\mathbf{w}_T^{(1)})^T \ (\mathbf{w}_T^{(2)})^T \ \cdots \ (\mathbf{w}_T^{(M)})^T\right]^T$, the output of the transform domain array system becomes

$$y_T(l) = \mathbf{w}_T^T \mathbf{x}_T(l) = \mathbf{w}_T^T \mathbf{T}\mathbf{x}(l). \tag{17}$$

Again, using the MMSE creterion

$$\begin{aligned} &\min_{\mathbf{w}_T} E\left|y_T(l) - s_1(l-v)\right|^2 \\ &= \min_{\mathbf{w}_T} E\left|\mathbf{w}_T^T \mathbf{x}_T(l) - s_1(l-v)\right|^2, \end{aligned} \tag{18}$$

the optimum weight vector becomes

$$\mathbf{w}_{T,opt} = \mathbf{R}_T^{-1}\mathbf{r}_T = (\mathbf{T}^T)^{-1}\mathbf{w}_{opt} \tag{19}$$

where

$$\mathbf{R}_T = E[\mathbf{x}_T^*(l)\mathbf{x}_T^T(l)] = \mathbf{T}^*\mathbf{R}\mathbf{T}^T \tag{20}$$

is the correlation matrix of $\mathbf{x}_T(l)$, and

$$\mathbf{r}_T = E[\mathbf{x}_T^*(l)s_1(l-v)] = \mathbf{T}^*\mathbf{r} \tag{21}$$

is the cross-correlation vector between $\mathbf{x}_T(l)$ and $s_1(l-v)$. When the optimum weight vectors are used for both STAP and the subband array, it is straightforward to show

$$y_T(l) = \mathbf{w}_{T,opt}^T \mathbf{T}\mathbf{x}(l) = \mathbf{w}_{opt}^T \mathbf{x}(l) = y(l), \tag{22}$$

and that the MSE of the subband array equals to the MMSE of the STAP systems

$$\begin{aligned} \text{MSE}_{CF} &= E\left|y_T(l) - s_1(l-v)\right|^2 \\ &= E\left|y(l) - s_1(l-v)\right|^2 \\ &= \text{MMSE}. \end{aligned} \tag{23}$$

IV. Subband Array with Localized Feedback

A. Structure

Subband arrays with the localized feedback scheme are often used for reduced system complexity and improved convergence performance. The basic idea behind the localized feedback is that the signal correlaton between signals at different subbands are often small due to the decorrelation function of the subband decomposition. Therefore, the signals at different subbands can be processed separately. A subband array with localized feedback scheme is illustrated in Fig. 2.

In the localized feedback scheme, the reference signal is decomposed into its subband version

$$s_1^{(m)}(l-v) = \frac{1}{\sqrt{M}} \mathbf{T}_o^{(m)} \vec{s}_1(l-v), \tag{24}$$

which is then used as the reference signal at the mth subband, where

$$\mathbf{T}_o^{(m)} = \frac{1}{\sqrt{M}} [W_M^0 \; W_M^m \; \cdots \; W_M^{(M-1)m}] \tag{25}$$

is the mth row of the matrix $\mathbf{T}_o$, and

$$\vec{s}_1(l-v) = [s_1(l-v) \; s_1(l-v-1) \; \cdots \; s_1(l-v-M+1)]^T$$

is the M samples of the reference signal used for the subband decomposition. The factor $1/\sqrt{M}$ used in (24) is to normalize the power of the reference signal at each subband because

$$\sum_{m=0}^{M-1} \mathbf{T}_o^{(m)} \vec{s}_1(l-v) = \left[\sum_{m=0}^{M-1} W_M^0 \sum_{m=0}^{M-1} W_M^0 \cdots \sum_{m=0}^{M-1} W_M^0 \right] \vec{s}_1(l-v) = \sqrt{M} s_1(l-v). \quad (26)$$

The $N \times 1$ weight vector at the mth subband, independent of other subbands, can be obtained from the $N \times N$ correlation matrix $\mathbf{R}_T^{(m)} = E\left[\mathbf{x}_T^{(m)}(l)(\mathbf{x}_T^{(m)}(l))^H\right]$ and the $N \times 1$ correlation vector $\mathbf{r}_T^{(m)} = E\left[\left(\mathbf{x}_T^{(m)}(l)\right)^* s_1^{(m)}(l-v)\right]$ as

$$\mathbf{w}_T'^{(m)} = (\mathbf{R}_T^{(m)})^{-1} \mathbf{r}_T^{(m)}. \quad (27)$$

B. Performance Analysis

Denote

$$\mathbf{R}_T' = \begin{bmatrix} \mathbf{R}_T^{(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_T^{(2)} & \cdots & \mathbf{0} \\ \vdots & & & \vdots \\ \mathbf{0} & \mathbf{0} & \vdots & \mathbf{R}_T^{(M)} \end{bmatrix} \quad (28)$$

and

$$\mathbf{r}_T' = \left[\left(\mathbf{r}_T^{(1)}\right)^T \quad \left(\mathbf{r}_T^{(2)}\right)^T \quad \cdots \quad \left(\mathbf{r}_T^{(M)}\right)^T \right]^T. \quad (29)$$

Using the following property of block-diagonal matrix

$$(\mathbf{R}_T')^{-1} = \begin{bmatrix} (\mathbf{R}_T^{(1)})^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{R}_T^{(2)})^{-1} & \cdots & \mathbf{0} \\ \vdots & & & \vdots \\ \mathbf{0} & \mathbf{0} & \vdots & (\mathbf{R}_T^{(M)})^{-1} \end{bmatrix}, \quad (30)$$

the weight vector of the localized feedback subband array can be expressed as

$$\mathbf{w}_T' = \begin{bmatrix} (\mathbf{R}_T^{(1)})^{-1}\mathbf{r}_T^{(1)} \\ (\mathbf{R}_T^{(2)})^{-1}\mathbf{r}_T^{(2)} \\ \vdots \\ (\mathbf{R}_T^{(M)})^{-1}\mathbf{r}_T^{(M)} \end{bmatrix} = (\mathbf{R}_T')^{-1}\mathbf{r}_T'. \quad (31)$$

As implied from (28), $\mathbf{R}'_T$ is the block-diagonal approximation of $\mathbf{R}_T$ by ignoring its off-block-diagonal elements. On the other hand, the cross-correlation vector between the received signal vector and the reference signal at the mth subband is

$$\begin{aligned}
\mathbf{r}_T^{(m)} &= E\left[\left(\mathbf{x}_T^{(m)}(l)\right)^* s_1^{(m)}(l-v)\right] \\
&= E\left[\left(\mathbf{T}^{(m)}\mathbf{x}(l)\right)^* s_1^{(m)}(l-v)\right] \\
&= E\left[\left(\mathbf{T}^{(m)}\right)^* \left(\sum_{p=1}^{P}\mathbf{H}_p\mathbf{s}_p(l)+\mathbf{b}(l)\right)^* \frac{1}{\sqrt{M}}\mathbf{T}_o^{(m)}\vec{s}_1(l-v)\right] \\
&= \frac{1}{\sqrt{M}}\left[\mathbf{T}^{(m)}\mathbf{H}_1\right]^* E\left[\mathbf{s}_1^*(l)\vec{s}_1^T(l-v)\right]\left[\mathbf{T}_o^{(m)}\right]^T \\
&= \frac{1}{\sqrt{M}}\left[\mathbf{T}^{(m)}\mathbf{H}_1\right]^* \mathbf{J}_v\left[\mathbf{T}_o^{(m)}\right]^T,
\end{aligned} \tag{32}$$

where $\mathbf{T}^{(m)}$ is the $N\times MN$ submatrix of the matrix $\mathbf{T}$ corresponding to the mth subband, $\mathbf{J}_v$ is an $(M+D_1-1)\times M$ matrix expressed as, provided that we choose $v<D_1$,

$$\mathbf{J}_v = E\left[\mathbf{s}_1^*(l)\vec{s}_1^T(l-v)\right] = [\mathbf{0}_v^T \;\; \mathbf{I}_M \;\; \mathbf{0}_{D_1-1-v}^T]^T, \tag{33}$$

where $\mathbf{0}_v$ denotes the zero matrix of size $v\times M$.

Therefore, the MSE of the localized feedback subband array is given by

$$\begin{aligned}
\mathrm{MSE}_{LF} &= E\left|s_1(l)-\mathbf{w}_T'^T\mathbf{x}_T(l)\right|^2 \\
&= 1+\mathbf{r}_T'^H(\mathbf{R}'_T)^{-1}\mathbf{R}_T(\mathbf{R}'_T)^{-1}\mathbf{r}'_T \\
&\quad -2\mathrm{Re}\left[\mathbf{r}_T'^H(\mathbf{R}'_T)^{-1}\mathbf{r}_T\right].
\end{aligned} \tag{34}$$

Equation (34) implies that the localized feedback subband array approach is suboptimal, and, its performance depends on the significance of the cross-correlation between signals at different subbands. It is clear from (20) and (34) that the off-block-diagonal elements of matrix $\mathbf{R}_T$ depends on both the transform matrix $\mathbf{T}$ and the channels $\mathbf{H}_p, p=1,2,...,P$.

V. Partial Feedback Scheme of Subband Arrays

In the previous section, we discussed the subband array with the localized feedback scheme as an approximation of the subband array with the centralized feedback scheme.

The former scheme has an independent weight update loop at each subband, at the cost of performance degradation, since the cross-correlations between different subbands are neglected in the weight estimation.

To provide more flexibility in trading-off the system performance and the complexity, we introduce subband arrays with the partial feedback scheme. As will be depicted, the partial feedback scheme is indeed a generalization of the centralized and localized feedback schemes, both can be considered as two extreme cases of the partial feedback scheme.

A subband array with partial feedback scheme is shown in Fig. 3, where the total M subbands are devided into K groups. The number of subbands in kth group is $M_k, k = 1, 2, ..., K$, with $M_1 + M_2 + \cdots + M_K = M$. In this paper, we consider the simple case of $M_1 = M_2 = \cdots = M_K = M/K$.

In this case, the signal covariance matrix $\mathbf{R}_T$ is approximated by a new block-diagonal matrix $\mathbf{R}_T''$ with a *larger* block size $M_1 N$, expressed as

$$\mathbf{R}_T'' = \begin{bmatrix} \mathbf{R}_T^{(G_1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_T^{(G_2)} & \cdots & \mathbf{0} \\ \vdots & & & \vdots \\ \mathbf{0} & \mathbf{0} & \vdots & \mathbf{R}_T^{(G_K)} \end{bmatrix} \tag{35}$$

where

$$\mathbf{R}_T^{(G_k)} = \begin{bmatrix} (\mathbf{R}_T)_{(k-1)M_1N+1,(k-1)M_1N+1} & \cdots & (\mathbf{R}_T)_{(k-1)M_1N+1,kM_1N} \\ \vdots & & \vdots \\ (\mathbf{R}_T)_{kM_1N,(k-1)M_1N+1} & \cdots & (\mathbf{R}_T)_{kM_1N,kM_1N} \end{bmatrix} \tag{36}$$

and $(\mathbf{R}_T)_{i,j}$ is the (i,j)-th element of matrix $\mathbf{R}_T$. When $M_1 > 1$, since fewer off-block-diagonal elements are ignored in $\mathbf{R}_T''$ as compared with $\mathbf{R}_T'$, the partial feedback scheme should provide more accurate optimum weights estimation and subsequently better MSE results than those of the localized feedback scheme.

Similar to (30), we have

$$(\mathbf{R}_T'')^{-1} = \begin{bmatrix} (\mathbf{R}_T^{(G_1)})^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{R}_T^{(G_2)})^{-1} & \cdots & \mathbf{0} \\ \vdots & & & \vdots \\ \mathbf{0} & \mathbf{0} & \vdots & (\mathbf{R}_T^{(G_K)})^{-1} \end{bmatrix}. \tag{37}$$

Therefore, the weight vector in the partial feedback scheme is given by

$$\mathbf{w}_T'' = (\mathbf{R}_T'')^{-1}\mathbf{r}''_T = \begin{bmatrix} (\mathbf{R}_T^{(G_1)})^{-1}\mathbf{r}_T^{(G_1)} \\ (\mathbf{R}_T^{(G_2)})^{-1}\mathbf{r}_T^{(G_2)} \\ \vdots \\ (\mathbf{R}_T^{(G_K)})^{-1}\mathbf{r}_T^{(G_K)} \end{bmatrix} \tag{38}$$

where

$$\mathbf{r}_T^{(G_k)} = E\left[\left(\mathbf{x}_T^{(G_k)}(l)\right)^* s_1^{(G_k)}(l)\right], \tag{39}$$

$$\mathbf{r}_T'' = \left[\left(\mathbf{r}_T^{(G_1)}\right)^T \quad \cdots \quad \left(\mathbf{r}_T^{(G_K)}\right)^T\right]^T, \tag{40}$$

$s_1^{(G_k)}$ is the reference signal at the kth group, and

$$\mathbf{x}_T^{(G_k)}(l)\left[\left(\mathbf{x}_T^{((k-1)M_1+1)}(l)\right)^T \cdots \left(\mathbf{x}_T^{(kM_1)}(l)\right)^T\right]^T. \tag{41}$$

The MSE of the partial feedback subband array is therefore

$$\begin{aligned} \mathrm{MSE}_{PF} &= E\left|s_1(l) - \mathbf{w}''^T_T\mathbf{x}_T(l)\right|^2 \\ &= 1 + \mathbf{r}''^H_T(\mathbf{R}_T'')^{-1}\mathbf{R}_T(\mathbf{R}_T'')^{-1}\mathbf{r}_T'' \\ &\quad -2\mathrm{Re}\left[\mathbf{r}''^H_T(\mathbf{R}_T'')^{-1}\mathbf{r}_T\right]. \end{aligned} \tag{42}$$

VI. Simulation Results

A three-element linear array with half wavelength inter-element spacing is considered. Two user signals are illuminating the array (P=2), each has a maximum delay spread of 5 symbols ($D = D_1 = D_2 = 5$). Six multipaths are randomly generated for each user whose detailed parameters are given in Tables 1 and 2, respectively. The input signal-to-noise ratio (SNR) is 20 dB for both signals.

Fig. 4(a) and (b) show the magnitude of the correlation matrices $\mathbf{R}$ and $\mathbf{R}_T$, where M=8. In Fig. 4(a), -60 dB is used to represent zero values so as to avoid errors in decibel calculation. In Fig. 5(a) and (b), we show similar results for M=32. It is clear that, while the value of $\mathbf{R}$ for different taps would be large depending on the channel coefficients, the value of $\mathbf{R}_T$ between different subbands becomes much smaller. However, the cost is increased floor values of the correlation matrix. The sidelobe effect is reduced as the number of subbands increases, as evident when comparing Fig. 4 and Fig. 5. This reduction is responsible for improving the MSE performance and pusing it closer to the optimum MMSE.

Fig. 6 shows the MSE performance for different feedback schemes. The number of subbands M changes from 4 to 32, and the MSE performance at different values of M_1 are evaluated. The dashed line shows the asymptotical lower bound of the MSE as M increases towards infinity. It is shown in Fig. 6 that the difference between different feedback schemes is large when M is relatively small (M is 4 or 8 in this figure) and small for large value of M (M is 16 or 32). Therefore, the subband array with localized or partial feedback schemes can closely approach the optimum MMSE performance when increasing the number of subbands.

VII. Conclusion

We have analyzed the performance of subband arrays with different types of feedback schemes, and the expressions of the steady-state mean square error (MSE) have been derived. It has been shown that subband arrays with localized and partial feedback schemes are generally suboptimal, and their performance depends on the channel characteristics, the filter banks employed, and the number of subbands. The proposed partial feedback scheme generalizes the subband arrays with centralized and localized feedback schemes, and provides more flexibility in trading-off the system complexity with the MSE performance.

Acknowledgment

Y. Zhang and K. Yang would like to thank Dr. B. Komiyama and Dr. T. Ohira, ATR Adaptive Communications Research Laboratories, Japan, for their encouragement and helpful discussions. Y. Zhang also thanks the valuable discussions by Prof. W. Wong and Prof. Z. Luo, McMaster University, Canada.

References

[1] Y. Ogawa, M. Ohmiya, and K. Itoh, "An LMS adaptive array for multipath fading reduction," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, no. 1, pp. 17–23, Jan. 1987.

[2] R. Kohno, "Spatial and temporal communication theory using adaptive antenna array," *IEEE Personal Communications*, vol. 5, no. 1, pp. 28–35, Feb. 1998.

[3] Y. Doi, T. Ohgane, and E. Ogawa, "ISI and CCI canceller combining the adaptive array antennas and the Viterbi equalizer in a digital mobile radio," in *Proc. IEEE VTC*, pp. 81–85, April 1996.

[4] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 49–83, Nov. 1997.

[5] Y. Zhang, K. Yang, and M. G. Amin, "Adaptive subband arrays for multipath fading mitigation," in *Proc. IEEE Antennas and Propagation Society Int. Symp.*, Atlanta, GA, pp. 380–383, June 1998.

[6] Y. Zhang, K. Yang, and M. G. Amin, "Performance analysis of subband adaptive arrays in multipath propagation environment," in *Proc. IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, Portland, OR, pp. 17–20, Sept. 1998.

[7] Y. Zhang, K. Yang, and Y. Karasawa, "Subband CMA adaptive arrays in multipath fading environment," *IEICE Trans. Commun.*, vol. J82-B, no. 1, pp. 97–108, Jan. 1999.

[8] Y. Zhang, K. Yang, and M. G. Amin, "Adaptive array processing for multipath fading mitigation via exploitation of filter banks," *IEEE Trans. Antennas Propagat.*, vol. 49, no. 4, April 2001.

[9] K. Yang, Y. Zhang, and Y. Mizuguchi, "Subband realization of space-time adaptive processing for mobile communications," in *Proc. Int. Symp. on Personal, Indoor and Mobile Radio Communications*, Osaka, Sept. 1999.

[10] K. Yang, Y. Zhang, and Y. Mizuguchi, "A signal subspace-based approach to space-time adaptive processing for mobile communications," *IEEE Trans. Signal Processing*, vol. 49, no. 2, pp. 401–413, Feb. 2001.

[11] J. M. Khalab and M. K. Ibrahim, "Novel multirate adaptive beamforming technique," *Electron. Lett.*, vol. 30. no. 15, pp. 1194-1195, 1994.

[12] Y. Zhang, K. Yang, and M. G. Amin, "Convergence performance of subband arrays for spatio-temporal equalization," in *Proc. IEEE Statistical Signal Processing Workshop*, Singapore, Aug. 2001.

[13] T. Sekiguchi and Y. Karasawa, "CMA adaptive array antennas using analysis and synthesis filter banks," *IEICE Trans. Fund.*, vol. E81-A, no. 8, pp. 1570–1577, Aug. 1998.

[14] H. Hoffman and S. Kogon, "Subband STAP in wideband radar systems," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, Cambridge, MA, pp. 256–260, March 2000.

[15] A. Steinhardt, "Subband STAP processing: the fifth generation," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, Cambridge, MA, pp. 1–6, March 2000.

[16] R. T. Compton, "The relationship between tapped delay-line and FFT processing in adaptive arrays," *IEEE Trans. Antennas Propagat.*, vol. 36, no. 1, pp. 15–26, Jan. 1988.

[17] Y. Kamiya and Y. Karasawa, "Performance comparison and improvement in adaptive arrays based on time and frequency domain signal processing," *IEICE Trans. Commun.*, vol. J82-A, no. 6, pp. 867–874, June 1999.

[18] F. Beaufays, "Transform-domain adaptive filters: an analytical approach," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 422-431, Feb. 1995.

[19] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley, MA: Wellesley-Cambridge Press, 1996.

TABLE I

PARAMETERS OF THE SIGNAL OF USER 1

	h_1	AOA (deg)
$\tau = 0$	$0.7016 + j0.0000$	33.54
$\tau = T$	$0.1188 + j0.0570$	18.06
$\tau = 2T$	$-0.1353 + j0.3165$	38.26
$\tau = 3T$	$-0.2231 - j0.1808$	5.89
$\tau = 4T$	$0.1476 + j0.2898$	34.79
$\tau = 5T$	$-0.3106 - j0.2945$	30.78

TABLE II

PARAMETERS OF THE SIGNAL OF USER 2

	h_2	AOA (deg)
$\tau = 0$	$0.6787 + j0.0000$	47.77
$\tau = T$	$0.1561 - j0.0592$	54.82
$\tau = 2T$	$-0.2173 + j0.3342$	68.07
$\tau = 3T$	$-0.2801 + j0.1987$	55.60
$\tau = 4T$	$-0.1119 + j0.2950$	39.89
$\tau = 5T$	$-0.3122 + j0.1938$	44.11

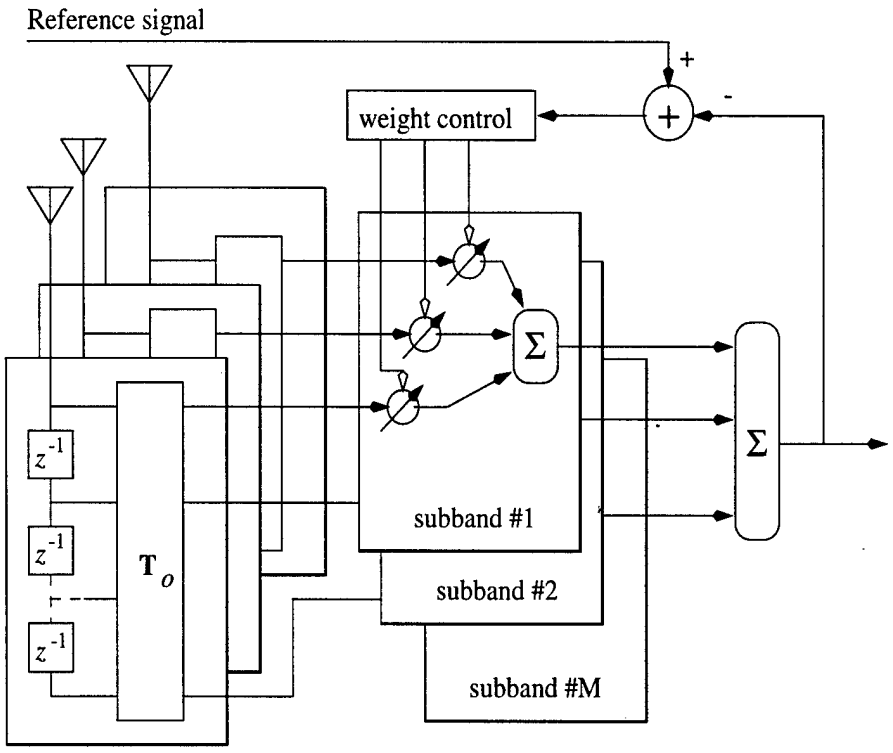


Fig. 1 Subband array with centralized feedback.

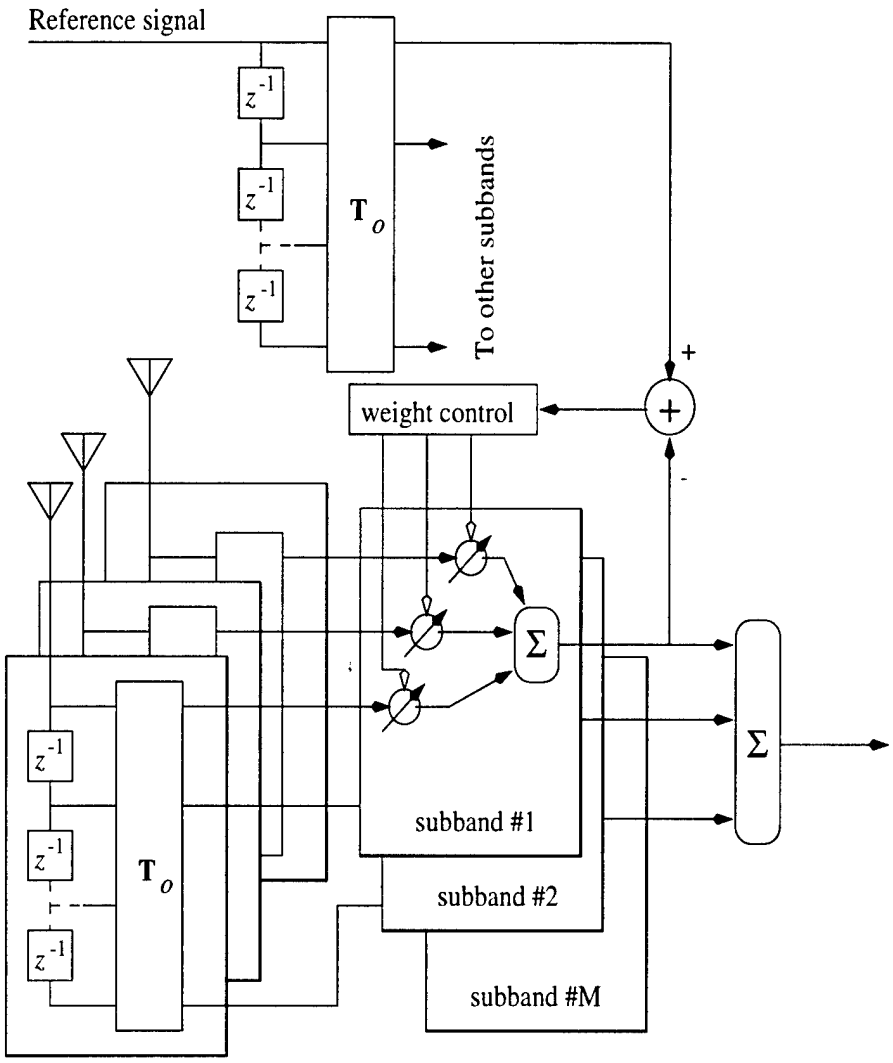


Fig. 2 Subband array with localozed feedback.

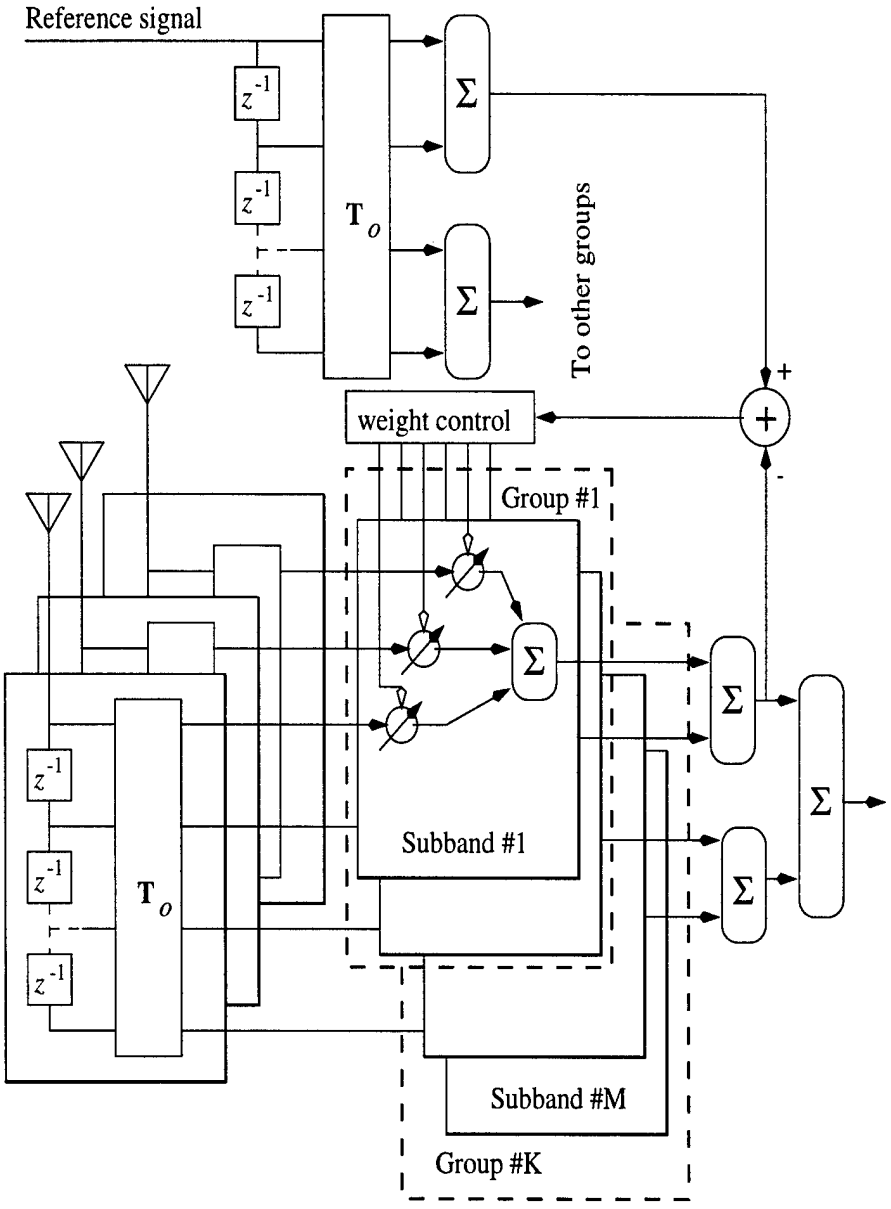
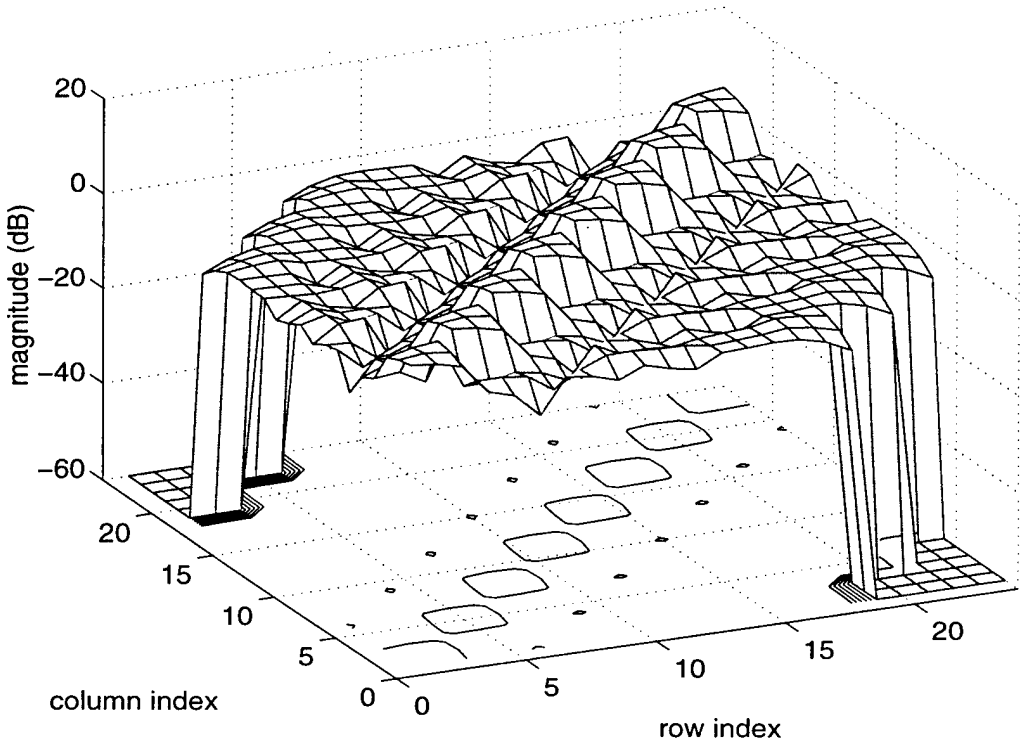
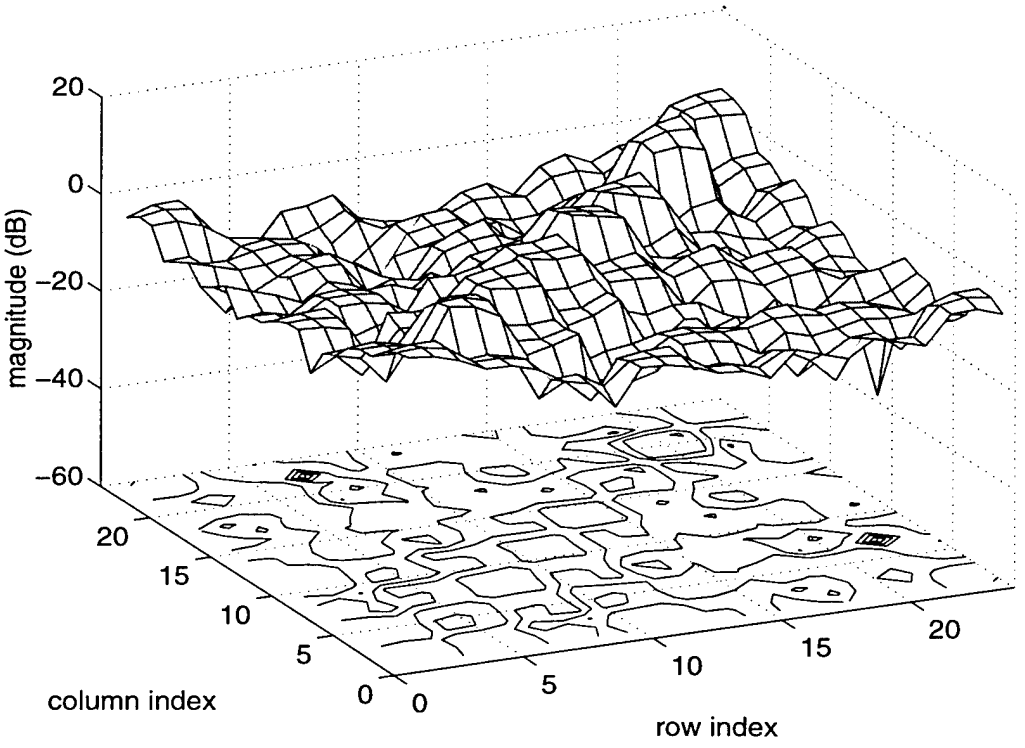


Fig. 3 Subband array with patial feedback.

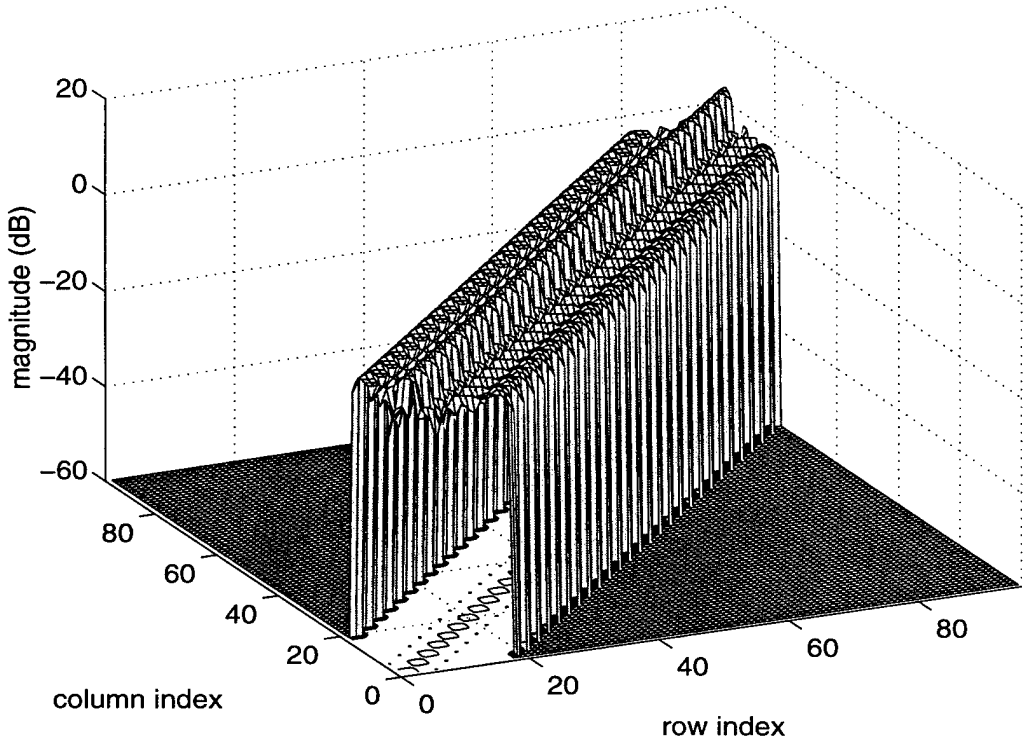


(a) $\mathbf{R}$

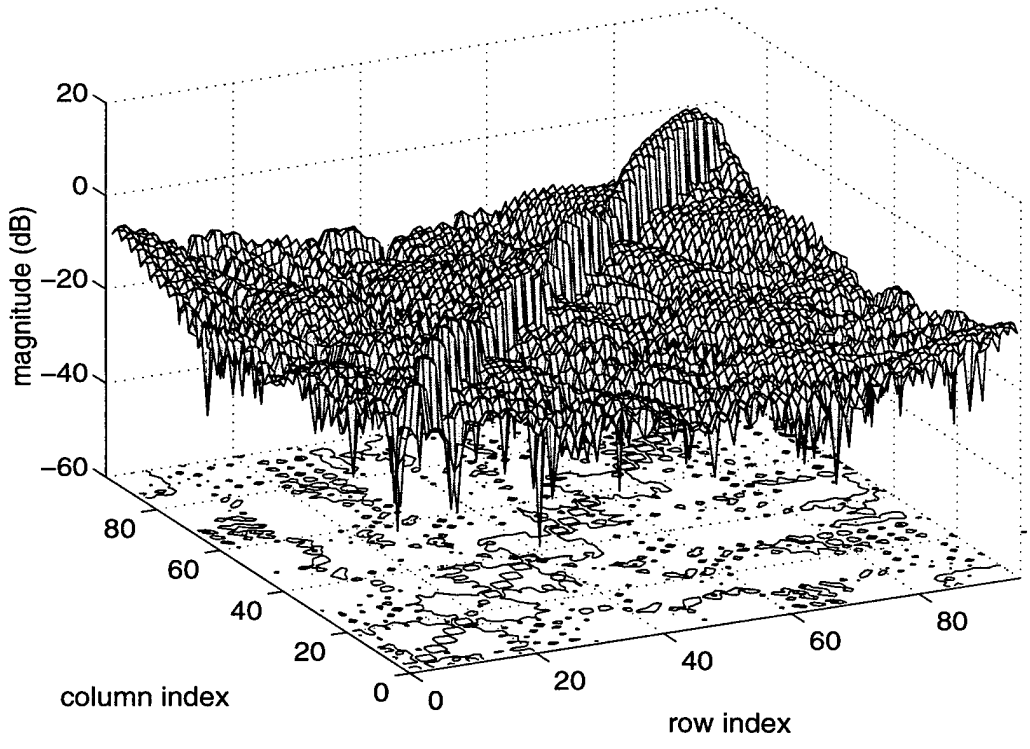


(b) $\mathbf{R}_T$

Fig. 4 Magnitudes of elements of $\mathbf{R}$ and $\mathbf{R}_T$ (M=8).



(a) $\mathbf{R}$



(b) $\mathbf{R}_T$

Fig. 5 Magnitudes of elements of $\mathbf{R}$ and $\mathbf{R}_T$ (M=32).

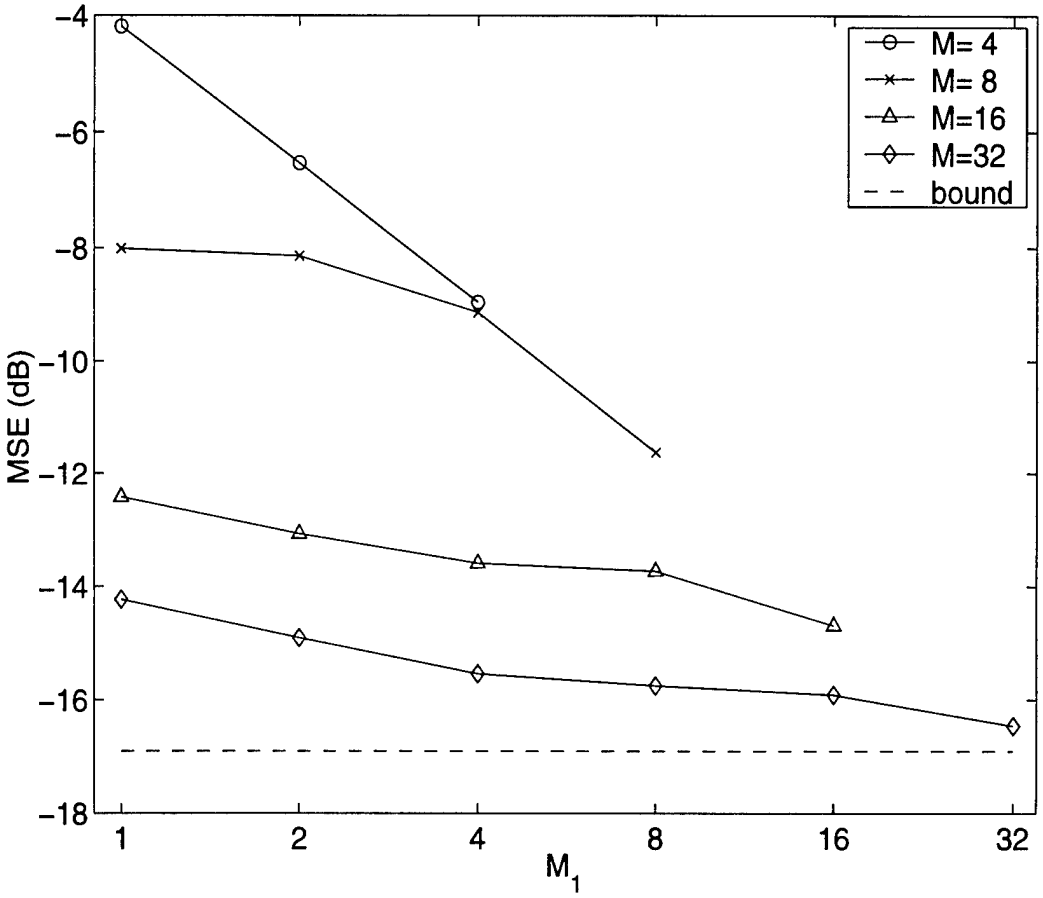


Fig. 6 MSE performance versus M and M_1.

Signal Separation based on Spatial Time Frequency Distributions Considered at Both Auto-Term and Cross-Term TF Points

A. Belouchrani[1], K. Abed-Meraim[2], M. G. Amin[3], A. M. Zoubir[4]
[1] Elec. Eng. Dept, Ecole Nationale Polytechnique,
EL Harrach 16200, Algiers, Algeria
[2] Sig. & Image Proc. Dept., Telecom Paris (ENST), Cedex 13, France
[3] Dept of Elec. and Comp. Eng.,
Villanova University, Villanova PA 19085 USA
[4] Australian Telecomm. Research Institute
and School of Elec. and Comp. Eng.,
Curtin University of Technology,
GPO Box U 1987, Perth 6485,

Abstract

We address the problem of blind source separation of non-stationary signals of which only instantaneous linear mixtures are observed. A blind source separation approach exploiting both auto-terms and cross-terms of the time-frequency (TF) distributions of the sources is considered. The approach is based on the simultaneous diagonalization and anti-diagonalization of a combined set of auto-term and cross-term time-frequency matrices, respectively. Numerical simulations are provided to demonstrate the effectiveness of our approach and compare its performances with existing TF-based methods.

THE WORK BY M. AMIN IS SPONSORED BY ONR, GRANT # N00014-98-1-0176.

I. Introduction

Blind source separation consists of recovering a set of signals of which only instantaneous linear mixtures are observed. Signal separation algorithms are based on the main assumption of mutual independence of the source signals. Various techniques have been proposed, including the separation by maximum likelihood [3], separation by decorrelation and rotation [1], [?], separation by neural networks [2], separation by contrast function [9], separation by information-theoretic criteria [10].

For non stationary source signals, blind source separation based on time-frequency distributions has been introduced in [5], [6]. The methods consider only auto-terms of the signal time-frequency distributions. They exploit the diagonal structure of the so-called Spatial Time-Frequency Distributions (STFDs) of the source signals, evaluated at the auto-term TF points. In this paper, we propose to exploit, in addition to the auto-terms, the cross-terms of the signal time-frequency distributions. This is achieved by exploiting the anti-diagonal structure of the STFDs of the source signals, evaluated at the cross-term TF points. Moreover, we propose an automatic selection procedure to decide, with no a priori knowledge about the sources, whether a considered TF point corresponds to an auto-term or a cross-term.

As a consequence, in comparison with the method in [5], the proposed one is more robust to noise and TF point selection errors and results in better separation quality of the sources.

II. Problem formulation

Consider m sensors receiving an instantaneous linear mixture of signals emitted from $n \leq m$ sources. The $m \times 1$ vector $\mathbf{x}(t)$ denotes the output of the sensors at time instant t which may be corrupted by additive noise $\mathbf{n}(t)$. Hence,

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \tag{1}$$

where the $m \times n$ matrix $\mathbf{A}$ is called the 'mixing matrix'. The n source signals are collected in a $n \times 1$ vector denoted $\mathbf{s}(t)$ which is referred to as the source signal vector. The sources are assumed to have different structures and localization properties in the time-frequency domain. The mixing matrix $\mathbf{A}$ is full column rank but is otherwise unknown. In contrast to traditional parametric methods, no specific structure of the mixture matrix is assumed.

The problem of blind source separation has two inherent ambiguities. First, it is not possible to know the original labeling of the sources, hence any permutation of the estimated sources is also a satisfactory solution. The second ambiguity is that it is inherently impossible to identify the scaling of the source signals. We take advantage of the second indeterminacy by treating the source signals as if they have *unit power*. This normalization still leaves undetermined the ordering and the phases of the columns of $\mathbf{A}$. Hence, the blind source separation is a technique for the identification of the mixing matrix and/or the recovering of the source signals up to a fixed permutation and some complex factors.

III. Spatial time-frequency distributions

The discrete-time form of Cohen's class of Time-Frequency Distributions (TFD) for a signal $x(t)$ is given by [7]

$$D_{xx}(t,f)=\sum_{l,m=-\infty}^{\infty}\phi(m,l)x(t+m+l)x^*(t+m-l)e^{-j4\pi fl} \tag{2}$$

where t and f represent the time index and the frequency index, respectively. The kernel $\phi(m,l)$ characterizes the distribution and is a function of both the time and lag variables. The cross-TFD of two signals $x_1(t)$ and $x_2(t)$ is defined by

$$D_{x_1x_2}(t,f)=\sum_{l,m=-\infty}^{\infty}\phi(m,l)x_1(t+m+l)x_2^*(t+m-l)e^{-j4\pi fl} \tag{3}$$

Expressions (2) and (3) are used to define the following data *spatial time-frequency distribution* (STFD) *matrix*,

$$\mathbf{D_{xx}}(t,f)=\sum_{l,m=-\infty}^{\infty}\phi(m,l)\mathbf{x}(t+m+l)\mathbf{x}^H(t+m-l)e^{-j4\pi fl} \tag{4}$$

where $[\mathbf{D_{xx}}(t,f)]_{ij} = D_{x_i x_j}(t,f)$, for $i,j = 1,\cdots,n$ and the superscript ^{H} denotes the transpose conjugate operator.

Under the linear data model (1) and assuming noise-free environment, the STFD matrix takes the following structure:

$$\mathbf{D_{xx}}(t,f) = \mathbf{A}\mathbf{D_{ss}}(t,f)\mathbf{A}^H \tag{5}$$

where $\mathbf{D_{ss}}(t,f)$ is the signal TFD matrix whose entries are the auto- and cross-TFDs of the sources.

Auto-STFD:. We define the auto-STFD by

$$\mathbf{D}^a_{\mathbf{ss}}(t,f) = \mathbf{D_{ss}}(t,f) \text{ for auto-term TF points} \tag{6}$$

Since the off-diagonal elements of $\mathbf{D_{ss}}(t,f)$ are cross-terms, the auto-STFD matrix is quasi diagonal for each TF point that corresponds to a true power concentration, i.e. signal auto-term.

Cross-STFD:. We define the Cross-STFD by

$$\mathbf{D}^c_{\mathbf{ss}}(t,f) = \mathbf{D_{ss}}(t,f) \text{ for cross-term TF points} \tag{7}$$

Since the diagonal elements of $\mathbf{D_{ss}}(t,f)$ are auto-terms, the cross-STFD matrix is quasi anti-diagonal (i.e. its diagonal entries are close to zero) for each TF point that corresponds to a cross-term.

IV. Proposed algorithm

Let $\mathbf{W}$ denote an $m \times n$ matrix, such that $(\mathbf{WA})(\mathbf{WA})^H = \mathbf{UU}^H = \mathbf{I}$, i.e. $\mathbf{WA}$ is an $m \times m$ unitary matrix ($\mathbf{W}$ is referred to as the whitening matrix, since it whitens the signal part of the observations). Pre- and post-multiplying the TFD-matrices $\mathbf{D_{xx}}(t,f)$ by $\mathbf{W}$, we define the *whitened TFD-matrices* as:

$$\underline{\mathbf{D}}_{\mathbf{xx}}(t,f) = \mathbf{W}\mathbf{D_{xx}}(t,f)\mathbf{W}^H \tag{8}$$

From the definition of $\mathbf{W}$ and Eq.(5), we can express $\underline{\mathbf{D}}_{\mathbf{xx}}(t,f)$ as

$$\underline{\mathbf{D}}_{\mathbf{xx}}(t,f) = \mathbf{U}\mathbf{D}_{\mathbf{ss}}(t,f)\mathbf{U}^H \tag{9}$$

Joint Diagonalization (JD):. By selecting auto-term TF points, the data auto-STFD will have the following structure,

$$\underline{\mathbf{D}}^a_{\mathbf{xx}}(t,f) = \mathbf{U}\mathbf{D}^a_{\mathbf{ss}}(t,f)\mathbf{U}^H \tag{10}$$

where $\mathbf{D}^a_{\mathbf{ss}}(t,f)$ is diagonal. The missing unitary matrix $\mathbf{U}$ is retrieved (up to permutation and phase shifts) by Joint Diagonalization (JD) of a combined set $\{\underline{\mathbf{D}}^a_{\mathbf{xx}}(t_i,f_i)|i=1,\cdots,p\}$ of p auto-STFD matrices. The incorporation of several auto-term TF points in the JD reduces the likelihood of having degenerate eigenvalues.

The joint diagonalization [4] of a set $\{\mathbf{M}_k|k=1..p\}$ of p $m\times m$ matrices is defined as the maximization of the JD criterion:

$$C(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{k=1}^{p}\sum_{i=1}^{m} |\mathbf{v}_i^H\mathbf{M}_k\mathbf{v}_i|^2 \tag{11}$$

over the set of unitary matrices $\mathbf{V}=[\mathbf{v}_1,\cdots,\mathbf{v}_m]$. An efficient joint approximate diagonalization algorithm exists in [4] and it is a generalization of the Jacobi technique [8] for the exact diagonalization of a single normal matrix.

Joint Anti-Diagonalization (JAD):. By selecting cross-term TF points, the data cross-STFD will have the following structure,

$$\underline{\mathbf{D}}^c_{\mathbf{xx}}(t,f) = \mathbf{U}\mathbf{D}^c_{\mathbf{ss}}(t,f)\mathbf{U}^H \tag{12}$$

where $\mathbf{D}^c_{\mathbf{ss}}(t,f)$ is anti-diagonal. The missing unitary matrix $\mathbf{U}$ is 'uniquely' (i.e. up to permutation and phase shifts) retrieved by Joint Anti-Diagonalization (JAD) of a combined set $\{\underline{\mathbf{D}}^c_{\mathbf{xx}}(t_i,f_i)|i=1,\cdots,q\}$ of q STFD matrices.

The joint anti-diagonalization is explained by first noting that the problem of anti-diagonalization of a single $m\times m$ matrix $\mathbf{N}$ is equivalent[1] to the maximization of the

[1]This is due to the fact that the Frobenius norm of a matrix is constant under unitary transform, i.e. norm($\mathbf{N}$) =norm($\mathbf{V}^H\mathbf{N}\mathbf{V}$).

criterion

$$C(\mathbf{N},\mathbf{V}) \stackrel{\text{def}}{=} -\sum_{i=1}^{m} |\mathbf{v}_i^H \mathbf{N} \mathbf{v}_i|^2 \tag{13}$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_m]$. Hence, JAD of a set $\{\mathbf{N}_k | k = 1..q\}$ of q $m \times m$ matrices is defined as the maximization of the JAD criterion:

$$C(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{k=1}^{q} C(\mathbf{N}_k,\mathbf{V}) = -\sum_{k=1}^{q}\sum_{i=1}^{m} |\mathbf{v}_i^H \mathbf{N}_k \mathbf{v}_i|^2 \tag{14}$$

under the same unitary constraint. A Jacobi-like algorithm has been derived for the maximization of the JAD criterion (14).

Combined JD/JAD algorithm:. The Combined joint diagonalization and joint anti-diagonalization of two sets $\{\mathbf{M}_k | k = 1..p\}$ and $\{\mathbf{N}_k | k = 1..q\}$ of $m \times m$ matrices is defined as the maximization of the JD/JAD criterion:

$$C(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{i=1}^{m} \left(\sum_{k=1}^{p} |\mathbf{v}_i^H \mathbf{M}_k \mathbf{v}_i|^2 - \sum_{k=1}^{q} |\mathbf{v}_i^H \mathbf{N}_k \mathbf{v}_i|^2 \right) \tag{15}$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_m]$. A Jacobi-like algorithm has been derived[2] for the maximization of the JD/JAD criterion (15).

Selection procedure:. The success of the JD or JAD of STFD matrices in determining the unitary matrix $\mathbf{U}$ depends strongly on the correct selection of the auto-term and cross-term points. A simulation example is given in Section 6. to emphasise this point. Therefore, it is crucial to have a selection procedure that is able to distinguish between auto-term and cross-term points based only on the STFD matrices of the observation. Here, we propose a selection approach that exploits the anti-diagonal structure of the cross-term STFD matrices. More precisely, we have

$$\begin{aligned} \text{Trace}(\underline{\mathbf{D}}_{\mathbf{xx}}^{c}(t,f)) &= \text{Trace}(\mathbf{U}\mathbf{D}_{\mathbf{ss}}^{c}(t,f)\mathbf{U}^H) \\ &= \text{Trace}(\mathbf{D}_{\mathbf{ss}}^{c}(t,f)) \approx 0. \end{aligned}$$

[2]Details of the JAD algorithm are omitted here due to space limitation.

Based on this observation, we derive the following testing procedure:

$$\text{if } \frac{\text{Trace}(\underline{\mathbf{D}}_{\mathbf{xx}}(t,f))}{\text{norm}(\underline{\mathbf{D}}_{\mathbf{xx}}(t,f))} < \epsilon \longrightarrow \text{ decide that } (t,f) \text{ is a cross-term}$$

$$\text{if } \frac{\text{Trace}(\underline{\mathbf{D}}_{\mathbf{xx}}(t,f))}{\text{norm}(\underline{\mathbf{D}}_{\mathbf{xx}}(t,f))} > \epsilon \longrightarrow \text{ decide that } (t,f) \text{ is an auto-term}$$

where ϵ is a 'small' positive real scalar. The correct choice of the value of ϵ is still under investigation. An ad-hoc value ($\epsilon = 0.1$) has been used in our simulation experiment.

Identification Procedure:. Equations (5-15) constitute the proposed blind source separation approach which is summarized by the following steps:

- Determine the whitening matrix $\hat{\mathbf{W}}$ from the eigen-decomposition of an estimate of the covariance matrix of the data (see [5] for more details).
- Compute the TF distribution of the array output according to (4).
- Select a set of TF points (usually corresponding to the high amplitude points of the signal TF transform) then distinguish between auto-term and cross-term points using the above selection procedure.
- Determine the unitary matrix $\hat{\mathbf{U}}$ by maximizing the JD/JAD criterion applied to the whitened STFD matrices computed at the selected TF points.
- Obtain an estimate of the mixture matrix $\hat{\mathbf{A}}$ as $\hat{\mathbf{A}} = \hat{\mathbf{W}}^{\#}\hat{\mathbf{U}}$, where the superscript $\#$ denotes the pseudo-inverse, and an estimate of the source signals $\hat{\mathbf{s}}(t)$ as $\hat{\mathbf{s}}(t) = \hat{\mathbf{U}}^{H}\mathbf{W}\mathbf{x}(t)$.

V. Discussion

We give here some comments to get more insight into the proposed blind source separation (BSS) method:

1) In practice, the source cross-STFD matrices will not be purely anti-diagonal. This is because some auto-terms, through their side lobes or main lobes, will intrude over the cross-term regions. The cross-terms will be however the dominant components. This situation is similar to the earlier work on joint digonalization of STFD selecting auto-term points [5], where the source auto-TFD matrix are not purely diagonal because of

cross-term intrusion. This impairment is taken care of thanks to the joint approximate character of the diagonalization and the anti-diagonalization algorithm and its robustness.

2) In contrast to the previously proposed Time-Frequency Separation (TFS) approach [5], the new proposed algorithm, allows selecting TF points in both auto-term and cross-term regions, as both regions provide separate key information about the signals. This results in a better separation performance (see simulation example in Section 6.).

3) The cross-term issues rise in both t-f and ambiguity domain based BSS. Therefore our blind separation method can be applied to both domains.

4) The smoothing kernel reduces the cross terms by re-distributing them across the t-f domain, rather than being concentrated at specific points where they can be confused with real energy. This re-distribution process will place some of these terms on the top of the autoterms, rendering the STFD matrix, constructed from autoterms, non-diagonal. So, in many cases, the Wigner-Ville distribution is more robust then any other distribution.

5) The JAD algorithm provides an estimate of the unitary matrix $\mathbf{U}$ and cross-STFD matrices $\mathbf{D}_{ss}^{c}(t,f)$. A necessary condition for the uniqueness of the solution is that the number of equations is greater than the total number of unknown parameters. This leads to the condition $q \leq m-1$ where q is the number of the $m \times m$ matrices to be fed to the JAD algorithm. Note that for the JD algorithm we need only $p \leq 1$ as necessary condition. A more detailed study on the identifiability of the problem will be given elsewhere.

VI. Simulation

First experiment:. We consider a uniform linear array of $n = 3$ sensors having half wavelength spacing and receiving signals from $m = 2$ sources in the presence of white Gaussian noise. The sources arrive from different directions $\phi_1 = 10$ and $\phi_2 = 20$ degrees. The

emitted signals are two chirps. The kernel used for the computation of the TFDs is the Wigner-Ville kernel. Eight STFD matrices are considered.

The performance is characterized in terms of signal rejection. The mean rejection level is defined as

$$\mathcal{I} \stackrel{\text{def}}{=} \sum_{q \neq p} E|\hat{\mathbf{A}}^{\#}\mathbf{A}_{pq}|^2 \tag{16}$$

We compare in Figure 1 the performance of the TFS algorithm proposed in [5] and the new algorithm for a signal-to-noise ratio (SNR) in the range [5 - 20 dB]. The mean rejection levels are evaluated here over 100 Monte Carlo runs with 1024 samples. It turns out that, in this case, the new algorithm performs slightly better than the TFS algorithm.

Second experiment:. In this experiment, we consider two chirp signals ($n = 2$), described by

$$\begin{aligned} s_1(t) &= \exp(-j0.004\pi t^2) \\ s_2(t) &= \exp(-j0.004\pi t^2 - j\pi 0.4t), \end{aligned}$$

embedded in noise leading to an SNR of 0 dB. We set $m = 5$. The Wigner-Ville distribution (WVD) of the mixture at the middle sensor is depicted in Figure 2. From Figure 2, we selected eight arbitrary TF points, among which one was a cross-term. Using the algorithm based on JD only, suggested in [5], we obtain the estimated signals, described by WVDs, shown in Figure 3. The figure clearly shows that the algorithm had failed. An estimate of the mean rejection rate was as high as 3 dB. However, if we apply the proposed method from Section 4., the results are more promising, leading to Figure 4 with a signal rejection level estimate of -26 dB. One may suggest to remove the cross-term, identified with the method suggested in Section 4., and run a JD algorithm based on the auto-terms only. The result of this approach is depicted in Figure 5. Although visually not noticeable, this approach is worse as the signal mean rejection level estimate is higher by circa 2 dB (-24dB).

Third experiment:. Here, we use three sources signals at 20 dB SNR. The number of antenna elements is again $m = 5$. The WVD of the mixture at the middle sensor is depicted in Figure 6. Six TF points are considered, among which six are cross-terms. The procedure described above was used with $\epsilon = 0.1$ to identify the auto-terms and cross-terms and the JD/JAD criterion ran.The result is depicted in Figure 7. It is clearly seen that the method performs very well with a rejection mean level estimate of -28 dB.

VII. Conclusions

In this paper, the problem of blind separation of linear spatial mixtures of non-stationary source signals based on time-frequency distributions has been investigated. A solution based on the hybrid diagonalization / anti-diagonalization of a combined set of spatial time-frequency distribution matrices, selected in both the auto-term and cross-term regions, has been proposed. Numerical simulations have been provided to illustrate the effectiveness of our approach.

References

[1] P. Comon, "Independent component analysis. a new concept," *Sig. Proc.*, vol. 36, pp. 287–314, 1994.

[2] A. Cichocki, W. Kasprzak, S. Amari, "Multi-layer neural networks with a local adaptive learning rule for blind separation of source signals," *Proc. Nolta, Las Vegas, USA*, Dec. 1995.

[3] A. Belouchrani, J.-F. Cardoso, "Maximum likelihood source separation by the EM technique: Deterministic and stochastic implementation," in *Proc. Nolta, Las Vegas, USA*, Dec. 1995.

[4] A. Belouchrani, K. Abed-Meraim, J.-F Cardoso, E. Moulines, "A blind source separation technique using second order statistics," *IEEE Trans. on SP*, pp. 434–444, Feb. 1997.

[5] A. Belouchrani, M. G. Amin, "Blind source separation based on time-frequency signal representation," *IEEE Trans. on SP*, pp. 2888–2898, Nov. 1998.

[6] Y. Zhang, M. G. Amin, "Signal averaging of time-frequency distributions for signal recovery in uniform linear arrays," *IEEE Trans. on SP*, pp. 2892–2902, Oct. 2000.

[7] L. Cohen, *Time-frequency analysis.* Prentice Hall, 1995.

[8] G.H. Golub, C.F. Van Loan, *Matrix computations.* The Johns Hopkins University Press, 1989.

[9] P. Comon and E. Moreau, "Improved contrast dedicated to blind separation in communications", *Proc ICASSP, Munich*, pp. 3453–3456, 1997.

[10] J.F. Cardoso, "Infomax and maximum likelihood for blind source separation", *IEEE SPL*, pp. 112–114, Apr. 1997.

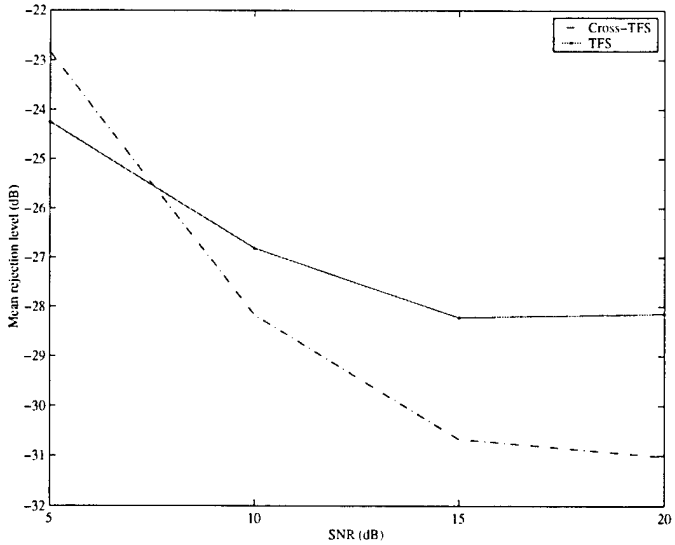


Fig. 1. Mean rejection level vs SNR.

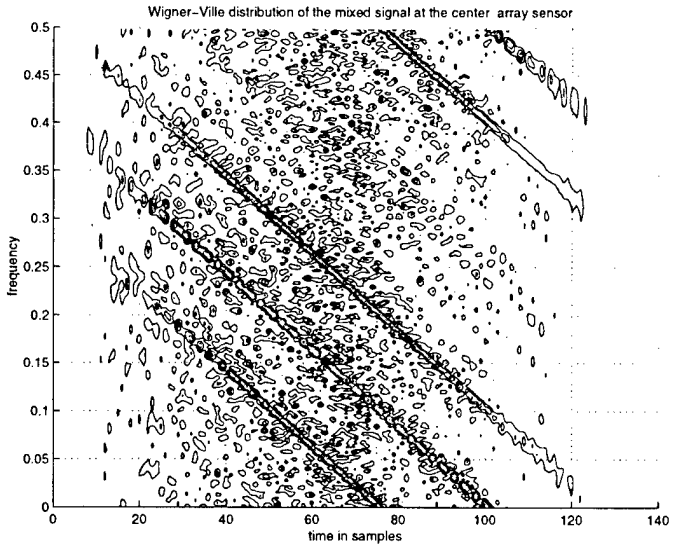


Fig. 2. WVD of two mixed signals at 0 dB SNR.

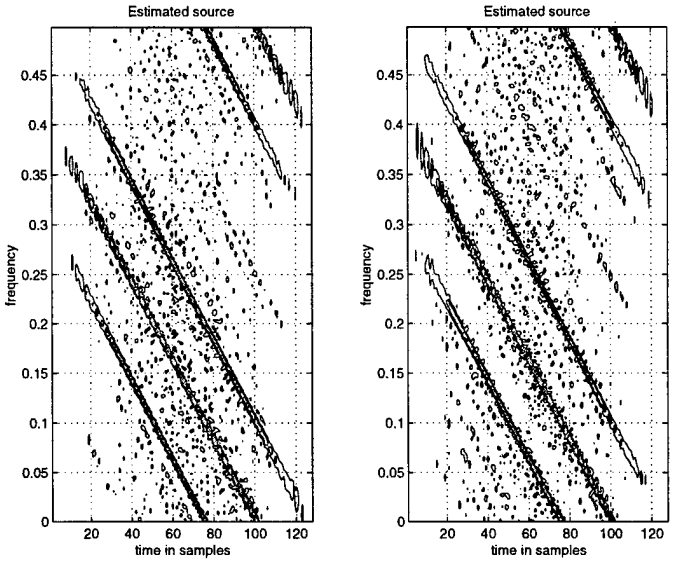


Fig. 3. WVDs of the two chirps using JD with seven auto-terms and one cross-term.

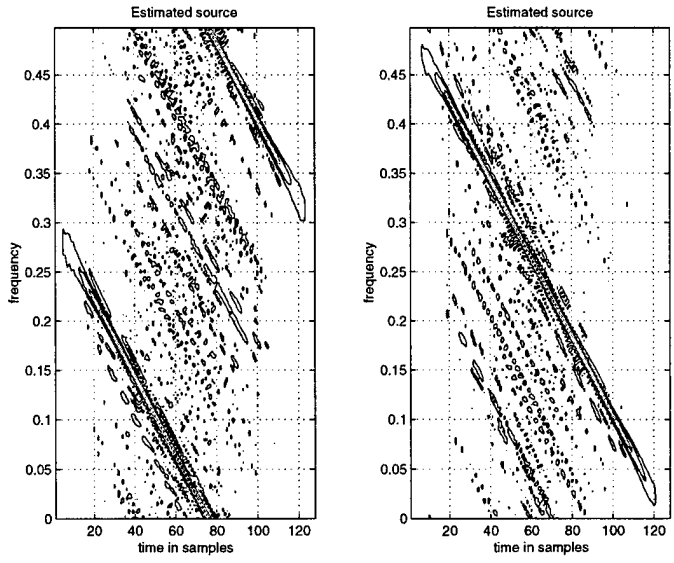


Fig. 4. WVDs of the two chirps using JD/JAD with seven auto-terms and one cross-term.

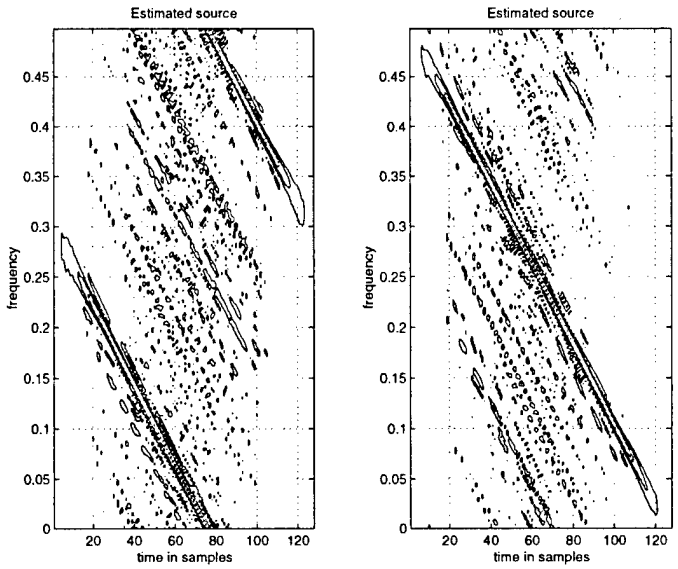


Fig. 5. WVDs of the two chirps using JD with the seven auto-terms from above.

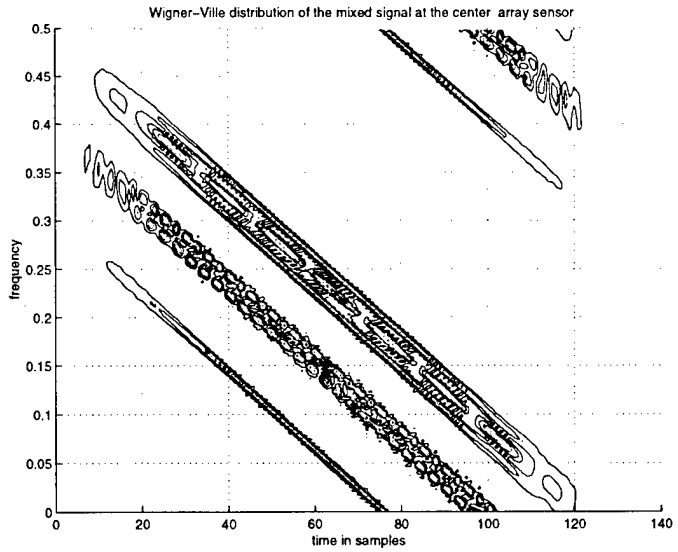


Fig. 6. WVD of the mixture of three chirps at 20 dB SNR.

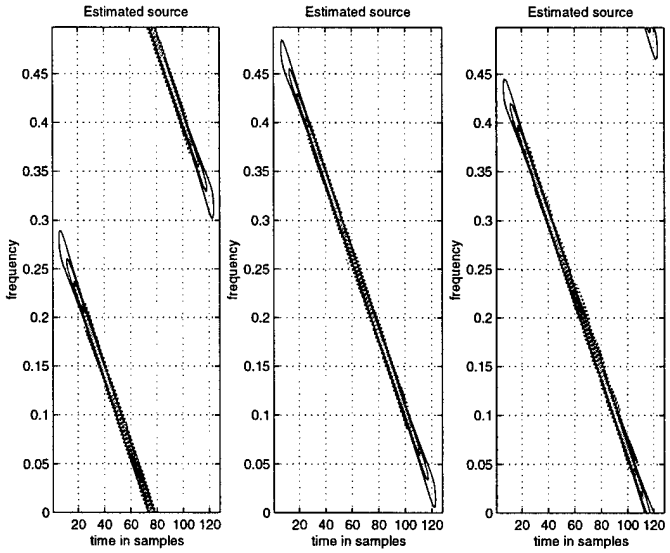


Fig. 7. WVDs of the three chirps using JD/JAD with one auto-term and five cross-terms.

Subspace Analysis of Spatial Time-Frequency Distribution Matrices

Yimin Zhang, Weifeng Mu, and Moeness G. Amin
Department of Electrical and Computer Engineering,
Villanova University, Villanova, PA 19085

Abstract

Spatial time-frequency distributions (STFDs) have been recently introduced as the natural means to deal with source signals that are localizable in the time-frequency domain. Previous work in the area has not provided the eigen analysis of STFD matrices, which is key to understanding their role in solving direction finding and blind source separation problems in multi-sensor array receivers. The aim of this paper is to examine the eigenstructure of the STFDs matrices. We develop the analysis and statistical properties of the subspace estimates based on STFDs for frequency modulated (FM) sources. It is shown that improved estimates are achieved by constructing the subspaces from the time-frequency signatures of the signal arrivals rather than from the data covariance matrices, which is commonly used in conventional subspace estimation methods. This improvement is evident in low signal-to-noise ratio (SNR) environment and in the cases of closely spaced sources. The paper considers the MUSIC technique to demonstrate the advantages of STFDs and uses it as grounds for comparison between time-frequency and conventional subspace estimates.

Keywords

Time-frequency distribution, subspace analysis, time-frequency MUSIC, spatial time-frequency distribution, array signal processing.

This work is supported by Office of Naval Research under Grant no. N00014-98-1-0176.

I. Introduction

While time-frequency distributions(TFDs) [1], [2], [3], [4] have been sought out and successfully used in the areas of speech, biomedicine, automotive industry, and machine monitoring, their use in sensor and spatial signal processing has not been properly investigated. The evaluation of quadratic time frequency distributions of the data snapshots across the array yields the "spatial time-frequency distributions" (STFDs) [5], [6]. These spatial distributions permit the application of eigenstructure subspace techniques to solving a large class of channel estimation and equalization, blind source separation, and high-resolution direction-of-arrival (DOA) estimation problems. In the area of blind source separation, the spatial time-frequency distributions allow the separation of Gaussian sources with identical spectral shape, but with different time-frequency localization properties, i.e., different signatures in the time-frequency domain. For direction-of-arrival estimation problems, the construction of the signal and noise subspaces using the source time-frequency signatures improves angular resolution performance.

Although the applications of the spatial time-frequency distributions to blind source separation and DOA estimation problems using multiple antenna arrays in nonstationary environments have been introduced in [5], [7], [8], yet so far there has not been sufficient analysis that explains their offerings and justify their performance. The aim of this paper is to examine the eigenstructure of the spatial time-frequency distribution matrices and provide statistical analysis of their respective signal and noise subspaces. The paper focuses on the class of frequency modulated (FM) signals, as they represent a clear case of nonstationary signals that are localizable in the time-frequency domain. It shows that the subspaces obtained from the STFDs are robust to both noise and angular separation of the FM waveforms incident on the array. This robustness is primarily due to spreading the noise power while localizing the source energy in the time-frequency domain. By forming the STFD matrices from the points residing on the source time-frequency signatures, we increase, in essence, the input signal-to-noise ratio (SNR), and hence improve the accuracy of the subspace estimates.

This paper is organized as follows. Section II presents the signal model and considers nonstationary environments defined by FM source signals. The statistical properties of

signal and noise subspace estimates for uncorrelated FM signals over the observation period are delineated. In Section III, we give a brief review of the definition and basic properties of the spatial time-frequency distributions, and derive the signal and noise subspaces using STFD matrices for the general class of FM signals. We demonstrate the robustness of the STFD-based subspace estimates to both noise and angular source separation as compared to those obtained in Section II using covariance matrices. The analytical results of Sections II and III are used in Section IV to examine the performance of the direction finding MUSIC technique based on the covariance matrix and STFD noise subspace estimates. Numerical simulations are given in Section V.

II. Subspace Analysis for FM Signals

A. Signal Model

In narrowband array processing, when n signals arrive at an m-element array, the linear data model

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{d}(t) + \mathbf{n}(t) \tag{1}$$

is commonly assumed, where the $m \times n$ spatial matrix $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$ represents the mixing matrix or the steering matrix. In direction finding problems, we require $\mathbf{A}$ to have a known structure, and each column of $\mathbf{A}$ corresponds to a single arrival and carries a clear bearing. On the other hand, when we consider blind source separation problems, $\mathbf{A}$ is a mixture of several steering vectors, due to multipaths, and its columns may assume any structure. The analytical treatment in this paper does not depend on any special structure of matrix $\mathbf{A}$.

Due to the mixture of the signals at each sensor, the elements of the $m \times 1$ data vector $\mathbf{x}(t)$ are multicomponent signals, whereas each source signal $d_i(t)$ of the $n \times 1$ signal vector $\mathbf{d}(t)$ is often a monocomponent signal. $\mathbf{n}(t)$ is an additive noise vector whose elements are modeled as stationary, spatially and temporally white, zero-mean complex random processes, independent of the source signals. That is,

$$E[\mathbf{n}(t+\tau)\mathbf{n}^H(t)] = \sigma\delta(\tau)\mathbf{I} \quad \text{and} \quad E[\mathbf{n}(t+\tau)\mathbf{n}^T(t)] = \mathbf{0} \quad \text{for any } \tau \tag{2}$$

where $\delta(\tau)$ is the delta function, $\mathbf{I}$ denotes the identity matrix, σ is the noise power at

each sensor, superscript ^{H} and ^{T}, respectively, denote conjugate transpose and transpose, and $E(\cdot)$ is the statistical expectation operator.

In equation (1), it is assumed that the number of sensors is larger than the number of sources, i.e., $m > n$. Further, matrix $\mathbf{A}$ is full column rank, which implies that the steering vectors corresponding to n different angles of arrival are linearly independent. We further assume that the correlation matrix

$$\mathbf{R}_{\mathbf{xx}} = E[\mathbf{x}(t)\mathbf{x}^H(t)] \tag{3}$$

is nonsingular, and the observation period consists of N snapshots with $N > m$.

Under the above assumptions, the correlation matrix is given by

$$\mathbf{R}_{\mathbf{xx}} = E[\mathbf{x}(t)\mathbf{x}^H(t)] = \mathbf{A}\mathbf{R}_{\mathbf{dd}}\mathbf{A}^H + \sigma\mathbf{I}, \tag{4}$$

where $\mathbf{R}_{\mathbf{dd}} = E[\mathbf{d}(t)\mathbf{d}^H(t)]$ is the source correlation matrix.

Let $\lambda_1 > \lambda_2 > \cdots > \lambda_n > \lambda_{n+1} = \lambda_{n+2} = \cdots = \lambda_m = \sigma$ denote the eigenvalues of $\mathbf{R}_{\mathbf{xx}}$. It is assumed that λ_i, $i = 1, \cdots, n$, are distinct. The unit-norm eigenvectors associated with $\lambda_1, ..., \lambda_n$ constitute the columns of matrix $\mathbf{S} = [\mathbf{s}_1, ..., \mathbf{s}_n]$, and those corresponding to $\lambda_{n+1}, ..., \lambda_m$ make up matrix $\mathbf{G} = [\mathbf{g}_1, ..., \mathbf{g}_{m-n}]$. Since the columns of $\mathbf{A}$ and $\mathbf{S}$ span the same subspace, then $\mathbf{A}^H\mathbf{G} = \mathbf{0}$.

In practice, $\mathbf{R}_{\mathbf{xx}}$ is unknown, and therefore should be estimated from the available data samples (snapshots) $\mathbf{x}(i)$, $i = 1, 2, ..., N$. The estimated correlation matrix is given by

$$\hat{\mathbf{R}}_{\mathbf{xx}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}(i)\mathbf{x}^H(i). \tag{5}$$

Let $\{\hat{\mathbf{s}}_1, ..., \hat{\mathbf{s}}_n, \hat{\mathbf{g}}_1, ..., \hat{\mathbf{g}}_{m-n}\}$ denote the unit-norm eigenvectors of $\hat{\mathbf{R}}_{\mathbf{xx}}$, arranged in the descending order of the associated eigenvalues, and let $\hat{\mathbf{S}}$ and $\hat{\mathbf{G}}$ denote the matrices defined by the set of vectors $\{\hat{\mathbf{s}}_i\}$ and $\{\hat{\mathbf{g}}_i\}$, respectively. The statistical properties of the eigenvectors of the sample covariance matrix $\hat{\mathbf{R}}_{\mathbf{xx}}$ for signals modeled as independent processes with additive white noise are given in [9].

B. Subspace Analysis for FM Signals

In this paper, we focus on analytic frequency modulation (FM) signals, modeled as

$$\mathbf{d}(t) = [d_1(t), ..., d_n(t)]^T = [D_1 e^{j\psi_1(t)}, ..., D_n e^{j\psi_n(t)}]^T, \tag{6}$$

where D_i and $\psi_i(t)$ are the fixed amplitude and time-varying phase of ith source signal. For each sampling time t, $d_i(t)$ has an instantaneous frequency $f_i(t) = \frac{1}{2\pi}\frac{d\psi_i(t)}{dt}$.

FM signals are often encountered in applications such as radar and sonar [2]. The consideration of FM signals in this paper is motivated by the fact that these signals are uniquely characterized by their instantaneous frequencies, and therefore, they have clear time-frequency signatures that can be utilized by the STFD approach. Also, FM signals have constant amplitudes. To simplify the analysis, we assume that the transmitted signals propagate in a stationary environment and are mutually uncorrelated over the observation period $[1 : N]$. Subsequently, the corresponding covariance matrices are time-independent. Under these assumptions, we have

$$\frac{1}{N}\sum_{k=1}^{N} d_i(k)d_j^*(k) = 0 \qquad \text{for } i \neq j,\ i,j = 1,...,n. \tag{7}$$

In this case, the signal correlation matrix in (4) is

$$\mathbf{R_{dd}} = diag\left[D_i^2,\ i = 1,2,...,n\right]$$

where $diag[\cdot]$ is the diagonal matrix formed with the elements of its vector valued arguments. From the above assumptions, we have the following Lemma.

Lemma 1: For uncorrelated FM signals with additive white Gaussian noise,

a) The estimation errors $(\hat{\mathbf{s}}_i - \mathbf{s}_i)$ are asymptotically (for large N) jointly Gaussian distributed with zero means and covariance matrices given by

$$\begin{aligned} &E\left[(\hat{\mathbf{s}}_i - \mathbf{s}_i)(\hat{\mathbf{s}}_j - \mathbf{s}_j)^H\right] \\ &= \frac{\sigma}{N}\left[\sum_{\substack{k=1\\k\neq i}}^{n} \frac{\lambda_i + \lambda_k - \sigma}{(\lambda_k - \lambda_i)^2}\mathbf{s}_k\mathbf{s}_k^H + \sum_{k=1}^{m-n}\frac{\lambda_i}{(\sigma - \lambda_i)^2}\mathbf{g}_k\mathbf{g}_k^H\right]\delta_{i,j} \triangleq \tilde{\mathbf{W}}_i\delta_{i,j}, \end{aligned} \tag{8}$$

$$E\left[(\hat{\mathbf{s}}_i - \mathbf{s}_i)(\hat{\mathbf{s}}_j - \mathbf{s}_j)^T\right] = -\frac{\sigma}{N}\frac{(\lambda_i + \lambda_j - \sigma)}{(\lambda_j - \lambda_i)^2}\mathbf{s}_j\mathbf{s}_i^T(1 - \delta_{i,j}) \triangleq \tilde{\mathbf{V}}_{i,j}. \tag{9}$$

b) The orthogonal projections of $\{\hat{\mathbf{g}}_i\}$ onto the column space of $\mathbf{S}$ are asymptotically (for large N) jointly Gaussian distributed with zero means and covariance matrices given by

$$E\left[\left(\mathbf{S}\mathbf{S}^H\hat{\mathbf{g}}_i\right)\left(\mathbf{S}\mathbf{S}^H\hat{\mathbf{g}}_j\right)^H\right] = \frac{\sigma}{N}\left[\sum_{k=1}^{n}\frac{\lambda_k}{(\sigma - \lambda_k)^2}\mathbf{s}_k\mathbf{s}_k^H\right]\delta_{i,j} \triangleq \frac{1}{N}\tilde{\mathbf{U}}\delta_{i,j}, \tag{10}$$

$$E\left[\left(\mathbf{S}\mathbf{S}^H\hat{\mathbf{g}}_i\right)\left(\mathbf{S}\mathbf{S}^H\hat{\mathbf{g}}_j\right)^T\right] = 0 \quad \text{for all } i,\ j. \tag{11}$$

The proof of part a) is given in Appendix A. The proof of part b) follows the same exact steps of the respective results derived in reference [9], and is not given here. Equations (8) and (9) hold strong similarity to those of [9]. The only difference is that the term $(\lambda_i\lambda_k)$ in equations (3.8a) and (3.8b) in [9] is replaced by $\sigma(\lambda_i + \lambda_k - \sigma)$ in (8) and (9), due to the uncorrelation property (7). Accordingly, for high input SNR ($\lambda_k \gg \sigma, k = 1, 2, ..., n$), the estimation error of $(\hat{\mathbf{s}}_i - \mathbf{s}_i)$ can be greatly reduced. From (8) and (9), each column of the signal subspace will be perfectly estimated when $\sigma = 0$. This is in contrast with the estimation error that would result under the same noise-free condition, if we use the temporally-independent signal characteristics considered in [9].

Equations (10) and (11) are identical to (3.9) and (3.10) derived in reference [9]. The reason of such identity is that, despite the difference in the signal eigenvectors in the two different scenarios, discussed in this paper and in reference [9], the signal subspaces in both cases are identical, and is spanned by the columns of matrix **A**. Accordingly, the projection of the estimated noise eigenvectors on the true signal subspace for either FM signals or white random processes yield equal results.

III. Subspace Analysis for STFD Matrices

The purpose of this section is to show that the signal and noise subspaces based on time-frequency distributions for nonstationary signals are more robust to noise than those obtained from conventional array processing.

A. Spatial Time-Frequency Distributions

We first review the definition and basic properties of the spatial time-frequency distributions (STFDs). STFDs based on Cohen's class of time-frequency distribution were introduced in [5] and its applications to direction finding and blind source separation have been discussed in [7], [8] and [5], respectively. In this paper, we consider the simplest member of Cohen's class, namely the pseudo Wigner-Ville distribution (PWVD) [1] and its respective spatial distribution. Only the time-frequency (t-f) points in the autoterm regions of PWVD are considered for STFD matrix construction. The autoterm region

refers to the time-frequency points along the true instantaneous frequency (IF) of each signal. The crossterms may intrude on the autoterms through the power in their main-lobes or/and sidelobes. This intrusion depends on the signal temporal structures and the window size. In this paper, however, we assume that the crossterms are negligible over the autoterm regions.

The discrete form of pseudo Wigner-Ville distribution of a signal $x(t)$, using a rectangular window of odd length L, is given by

$$D_{xx}(t,f) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} x(t+\tau)x^*(t-\tau)e^{-j4\pi f\tau}, \tag{12}$$

where * denotes complex conjugate. It should be noted that incorporating multiple time-frequency points, via time-frequency averaging, over the autoterm region causes the crossterm components present at the signal IF to cancel each other, rendering their overall effect negligible.

The spatial pseudo Wigner-Ville distribution (SPWVD) matrix is obtained by replacing $x(t)$ by the data snapshot vector $\mathbf{x}(t)$,

$$\mathbf{D_{xx}}(t,f) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} \mathbf{x}(t+\tau)\mathbf{x}^H(t-\tau)e^{-j4\pi f\tau}. \tag{13}$$

Substitute (1) into (13), we obtain

$$\mathbf{D_{xx}}(t,f) = \mathbf{D_{yy}}(t,f) + \mathbf{D_{yn}}(t,f) + \mathbf{D_{ny}}(t,f) + \mathbf{D_{nn}}(t,f). \tag{14}$$

We note that $\mathbf{D_{xx}}(t,f)$, $\mathbf{D_{yy}}(t,f)$, $\mathbf{D_{yn}}(t,f)$, $\mathbf{D_{ny}}(t,f)$, and $\mathbf{D_{nn}}(t,f)$ are matrices of dimension $m \times m$. Under the uncorrelated signal and noise assumption and the zero-mean noise property, the expectation of the crossterm STFD matrices between the signal and noise vectors is zero, i.e., $E[\mathbf{D_{yn}}(t,f)] = E[\mathbf{D_{ny}}(t,f)] = \mathbf{0}$, and it follows

$$\begin{aligned} E[\mathbf{D_{xx}}(t,f)] &= \mathbf{D_{yy}}(t,f) + E[\mathbf{D_{nn}}(t,f)] \\ &= \mathbf{A}\mathbf{D_{dd}}(t,f)\mathbf{A}^H + E[\mathbf{D_{nn}}(t,f)], \end{aligned} \tag{15}$$

where the source time-frequency distribution (TFD) matrix

$$\mathbf{D_{dd}}(t,f) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} \mathbf{d}(t+\tau)\mathbf{d}^H(t-\tau)e^{-j4\pi f\tau} \tag{16}$$

is of dimension $n \times n$. For narrowband array signal processing applications, the mixing matrix $\mathbf{A}$ holds the spatial information and maps the auto- and cross-TFDs of the source signals into auto- and cross-TFDs of the data.

Equation (15) is similar to the formula that has been commonly used in DOA estimation and blind source separation problems, relating the signal correlation matrix to the data spatial correlation matrix. In the above formulation, however, the correlation matrices are replaced by the spatial time-frequency distribution matrices. The well established results in conventional array signal processing could, therefore, be utilized and key problems in various applications of array processing, specifically those dealing with nonstationary signal environments, can be approached using bilinear transformations.

It is noted that relationship (15) holds true for every (t, f) point. In order to reduce the effect of noise and ensure the full column rank property of the STFD matrix, we consider multiple time-frequency points, instead of a single one. That is, the signal and noise subspaces are constructed using as many (t, f) points in the source autoterm regions as possible. This allows more information of the source signal time-frequency signatures to be included into their respective subspace formulation, and as such enhances direction finding and source separation performance. Joint-diagonalization [10] and time-frequency averaging are the two main approaches that have been used for this purpose [5], [7], [11]. In this paper, however, we only consider averaging over multiple time-frequency points.

B. SNR Enhancement

The TFD maps one-dimensional signals in the time domain into two-dimensional signals in the time-frequency domain. The TFD property of concentrating the input signal around its instantaneous frequency while spreading the noise over the entire time-frequency domain increases the effective SNR and proves valuable in the underlying problem.

The ith diagonal element of TFD matrix $\mathbf{D_{dd}}(t, f)$ is given by

$$D_{d_i d_i}(t, f) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} D_i^2 e^{j[\psi_i(t+\tau)-\psi_i(t-\tau)]-j4\pi f\tau}. \tag{17}$$

Assume that the third-order derivative of the phase is negligible over the window length L, then along the true time-frequency points of the ith signal, $f_i(t) = \dfrac{1}{2\pi}\dfrac{d\psi_i(t)}{dt}$, and

$\psi_i(t+\tau) - \psi_i(t-\tau) - 4\pi f_i(t)\tau = 0$. Accordingly, for $\frac{L-1}{2} \leq t \leq N - \frac{L-1}{2}$,

$$D_{d_i d_i}(t, f_i(t)) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} D_i^2 = LD_i^2. \tag{18}$$

Similarly, the noise STFD matrix $\mathbf{D_{nn}}(t, f)$ is

$$\mathbf{D_{nn}}(t, f) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} \mathbf{n}(t+\tau)\mathbf{n}^H(t-\tau)e^{-j4\pi f\tau}. \tag{19}$$

Under the spatially and temporally white assumptions, the statistical expectation of $\mathbf{D_{nn}}(t, f)$ is given by

$$E\left[\mathbf{D_{nn}}(t, f)\right] = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} E\left[\mathbf{n}(t+\tau)\mathbf{n}^H(t-\tau)\right] e^{-j4\pi f\tau} = \sigma\mathbf{I}. \tag{20}$$

Therefore, when we select the time-frequency points along the time-frequency signature or the IF of the ith FM signal, the SNR in model (15) is LD_i^2/σ, which has an improved factor L over the one associated with model (4). The IF of the FM signals can be estimated from the employed TFD, which in this case is the PWVD. It may also be given separately using any appropriate IF estimator. It is noted that the STFD equation (15) provides a natural platform for the direct incorporation of any *a priori* information or estimates of the IF into DOA estimation.

The pseudo Wigner-Ville distribution of each FM source has a constant value over the observation period, providing that we leave out the rising and falling power distributions at both ends of the data record. For convenience of analysis, we select those $N - L + 1$ time-frequency points of constant distribution value for each source signal. In the case where the STFD matrices are averaged over the time-frequency signatures of n_o sources, i.e., a total of $n_o(N - L + 1)$ time-frequency points, the result is given by

$$\hat{\mathbf{D}} = \frac{1}{n_o(N-L+1)} \sum_{q=1}^{n_o} \sum_{i=1}^{N-L+1} \mathbf{D_{xx}}(t_i, f_{q,i}(t_i)), \tag{21}$$

where $f_{q,i}(t_i)$ is the instantaneous frequency of the qth signal at the ith time sample. $\mathbf{x}(t)$ is an instantaneous mixture of the FM signals $d_i(t), i = 1, \cdots, n$, hence features the same IFs. The expectation of the averaged STFD matrix is

$$\mathbf{D} = E\left[\hat{\mathbf{D}}\right] = \frac{1}{n_o(N-L+1)} \sum_{q=1}^{n_o} \sum_{i=1}^{N-L+1} E\left[\mathbf{D}_{\mathbf{xx}}(t_i, f_{q,i}(t_i))\right]$$

$$= \frac{1}{n_o} \sum_{q=1}^{n_o} \left[LD_q^2 \mathbf{a}_q \mathbf{a}_q^H + \sigma \mathbf{I}\right] = \frac{L}{n_o} \mathbf{A}^o \mathbf{R}_{\mathbf{dd}}^o (\mathbf{A}^o)^H + \sigma \mathbf{I}, \tag{22}$$

where $\mathbf{R}_{\mathbf{dd}}^o = diag\left[D_i^2, i = 1, 2, \cdots, n_o\right]$ and $\mathbf{A}^o = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_{n_o}]$ represent the signal correlation matrix and the mixing matrix formulated by considering n_o signals out of the total number of n signal arrivals, respectively.

It is clear from (22) that the SNR improvement $G = L/n_o$ (we assume $L > n_o$ throughout this paper) is inversely proportional to the number of sources contributing matrix $\mathbf{D}$. Therefore, from the SNR perspective, it is best to set $n_o = 1$, i.e., to select the sets of $N - L + 1$ time-frequency points that belong to individual signals one set at a time, and then separately evaluate the respective STFD matrices.

This procedure is made possible by the fact that STFD-based direction finding is, in essence, a discriminatory technique in the sense that it does not require simultaneous localization and extraction of all unknown signals received by the array. With STFDs, direction finding can be performed using STFDs of a subclass of the impinging signals with specific time-frequency signatures. In this respect, the proposed direction finding technique acts as a spatial filter, removing all other signals from consideration and, subsequently, saves any downstream processing that is required to separate interference and signals of interest. It is also important to note that with the ability to construct the STFD matrix from one or few signal arrivals, the well known $m > n$ condition on source localization using arrays can be relaxed, i.e., we can perform direction finding or source separation with the number of array sensors smaller than the number of impinging signals [6]. From the angular resolution perspective, closely spaced sources with different time-frequency signatures can be resolved by constructing two separate STFDs, each corresponds to one source, and then proceed with subspace decomposition for each STFD matrix, followed by an appropriate source localization method (MUSIC, for example). The drawback of performing several direction finding using different STFD matrices is of course the need for repeated computations of eigen-decompositions and source localizations.

C. Signal and Noise Subspaces Using STFDs

The following Lemma provides the relationship between the eigen-decompositions of the STFD matrices and the data covariance matrices used in conventional array processing.

Lemma 2: Let $\lambda_1^o > \lambda_2^o > \cdots > \lambda_{n_o}^o > \lambda_{n_o+1}^o = \lambda_{n_o+2}^o = \cdots = \lambda_m^o = \sigma$ denote the eigenvalues of $\mathbf{R}_{\mathbf{xx}}^o = \mathbf{A}^o \mathbf{R}_{\mathbf{dd}}^o (\mathbf{A}^o)^H + \sigma \mathbf{I}$, which is defined from a data record of a mixture of the n_o selected FM signals. Denote the unit-norm eigenvectors associated with $\lambda_1^o, ..., \lambda_{n_o}^o$ by the columns of $\mathbf{S}^o = [\mathbf{s}_1^o, ..., \mathbf{s}_{n_o}^o]$, and those corresponding to $\lambda_{n_o+1}^o, ..., \lambda_m^o$ by the columns of $\mathbf{G}^o = [\mathbf{g}_1^o, ..., \mathbf{g}_{m-n_o}^o]$. We also denote $\lambda_1^{tf} > \lambda_2^{tf} > \cdots > \lambda_{n_o}^{tf} > \lambda_{n_o+1}^{tf} = \lambda_{n_o+2}^{tf} = \cdots = \lambda_m^{tf} = \sigma^{tf}$ as the eigenvalues of $\mathbf{D}$ defined in (22). The superscript ^{tf} denotes that the associated term is derived from the STFD matrix $\mathbf{D}$. The unit-norm eigenvectors associated with $\lambda_1^{tf}, ..., \lambda_{n_o}^{tf}$ are represented by the columns of $\mathbf{S}^{tf} = [\mathbf{s}_1^{tf}, ..., \mathbf{s}_{n_o}^{tf}]$, and those corresponding to $\lambda_{n_o+1}^{tf}, ..., \lambda_m^{tf}$ are represented by the columns of $\mathbf{G}^{tf} = [\mathbf{g}_1^{tf}, ..., \mathbf{g}_{m-n_o}^{tf}]$. Then,

a) The signal and noise subspaces of $\mathbf{S}^{tf}$ and $\mathbf{G}^{tf}$ are the same as $\mathbf{S}^o$ and $\mathbf{G}^o$, respectively.

b) The eigenvalues have the following relationship:

$$\lambda_i^{tf} = \begin{cases} \dfrac{L}{n_o}(\lambda_i^o - \sigma) + \sigma = \dfrac{L}{n_o}\lambda_i^o + \left(1 - \dfrac{L}{n_o}\right)\sigma & i \le n_o \\ \sigma^{tf} = \sigma & n_o < i \le m. \end{cases} \tag{23}$$

The proof of Lemma 2 is shown in Appendix B.

An important conclusion from Lemma 2 is that, the largest n_o eigenvalues are amplified using STFD analysis. Fig. 1 shows the two principal (largest) eigenvalues λ_i^o (i.e., $L = 1$) and λ_i^{tf} (for $L = 33$ and $L = 129$), where a uniform linear array of eight sensors ($m = 8$) separated by half a wavelength and receiving signal from two sources ($n_o = n = 2$) is used. The two signals are of equal power($D_1 = D_2 = D$), and their angular separation $\Delta\theta$ is defined as $\theta_2 - \theta_1$. We choose $\theta_1 + \theta_2 = 0$, that is, the two signals are symmetric with respect to the broadside direction. Denote

$$\beta = \frac{\mathbf{a}_1^H \mathbf{a}_2}{||\mathbf{a}_1||_2 \, ||\mathbf{a}_2||_2}$$

as the spatial correlation coefficient between the two directional vectors $\mathbf{a}_1$ and $\mathbf{a}_2$, corresponding to the angles θ_1 and θ_2. $||\mathbf{a}||_2$ is the 2-norm of a vector $\mathbf{a}$. The two largest

eigenvalues for the two uncorrelated signals are given by [12]

$$\lambda_{1,2}^{o} = mD^2 \left[1 \pm |\beta|\right] + \sigma. \tag{24}$$

Hence, combining (23) and (24), we obtain

$$\lambda_{1,2}^{tf} = \frac{mL}{n_o} D^2 \left[1 \pm |\beta|\right] + \sigma. \tag{25}$$

The amplification of the largest n_o eigenvalues improves detection of the number of the impinging signals on the array, as it widens the separation between dominant and noise-level eigenvalues. Determination of the number of signals is key to establishing the proper signal and noise subspaces, and subsequently plays a fundamental role in subspace-based applications [13]. When the input SNR is low, or the signals are closely spaced, the number of signals may often be underdetermined. Fig. 2 shows, for the same signal scenario of Fig. 1, the threshold level of the input SNR required to determine the correct number of signals $\hat{n} = 2$ according to the Akaike Information Criterion (AIC) [14]

$$\min_{\hat{n}} N(m - \hat{n}) log \left[\frac{f_1(\hat{n})}{f_2(\hat{n})}\right] + f_3(\hat{n}), \tag{26}$$

where

$$f_1(\hat{n}) \triangleq \frac{1}{m - \hat{n}} \sum_{i=\hat{n}+1}^{m} \lambda_i, \quad f_2(\hat{n}) \triangleq \left(\prod_{i=\hat{n}+1}^{m} \lambda_i \right)^{\frac{1}{m-\hat{n}}}, \quad f_3(\hat{n}) \triangleq \hat{n}(2m - \hat{n}). \tag{27}$$

It is clear from Fig. 2 that, when the STFD is applied, the SNR threshold level necessary for the correct determination of the number of signals is greatly reduced.

Next we consider the signal and noise subspace estimates from a finite number of data samples. We form the STFD matrix based on the true (t, f) points along the IF of the n_o FM signals.

Lemma 3: If the third-order derivative of the phase of the FM signals is negligible over the time-period $[t - L + 1, t + L - 1]$, then

a) The estimation errors in the signal vectors are asymptotically (for $N \gg L$) jointly Gaussian distributed with zero means and covariance matrices given by

$$E\left(\hat{\mathbf{s}}_i^{tf}-\mathbf{s}_i^{tf}\right)\left(\hat{\mathbf{s}}_j^{tf}-\mathbf{s}_j^{tf}\right)^H$$

$$= \frac{\sigma L}{n_o(N-L+1)}\left[\sum_{\substack{k=1\\k\neq i}}^{n_o}\frac{\lambda_i^{tf}+\lambda_k^{tf}-\sigma}{(\lambda_k^{tf}-\lambda_i^{tf})^2}\mathbf{s}_k^{tf}\left(\mathbf{s}_k^{tf}\right)^H+\sum_{k=1}^{m-n_o}\frac{\lambda_i^{tf}}{(\sigma-\lambda_i^{tf})^2}\mathbf{g}_k^{tf}\left(\mathbf{g}_k^{tf}\right)^H\right]\delta_{i,j}$$

$$\triangleq \mathbf{W}_i^{tf}\delta_{i,j}, \tag{28}$$

$$E\left(\hat{\mathbf{s}}_i^{tf}-\mathbf{s}_i^{tf}\right)\left(\hat{\mathbf{s}}_j^{tf}-\mathbf{s}_j^{tf}\right)^T$$

$$= -\frac{\sigma L}{n_o(N-L+1)}\frac{(\lambda_i^{tf}+\lambda_j^{tf}-\sigma)}{(\lambda_j^{tf}-\lambda_i^{tf})^2}\mathbf{s}_j^{tf}\left(\mathbf{s}_i^{tf}\right)^T(1-\delta_{i,j})\triangleq\mathbf{V}_{i,j}^{tf}. \tag{29}$$

b) The orthogonal projections of $\left\{\hat{\mathbf{g}}_i^{tf}\right\}$ onto the column space of $\mathbf{S}^{tf}$ are asymptotically (for $N\gg L$) jointly Gaussian distributed with zero means and covariance matrices given by

$$E\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H\hat{\mathbf{g}}_i^{tf}\right)\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H\hat{\mathbf{g}}_j^{tf}\right)^H$$

$$= \frac{\sigma L}{n_o(N-L+1)}\left[\sum_{k=1}^{n_o}\frac{\lambda_k^{tf}}{(\sigma-\lambda_k^{tf})^2}\mathbf{s}_k^{tf}\left(\mathbf{s}_k^{tf}\right)^H\right]\delta_{i,j}\triangleq\frac{1}{(N-L+1)}\mathbf{U}^{tf}\delta_{i,j}, \tag{30}$$

$$E\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H\hat{\mathbf{g}}_i^{tf}\right)\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H\hat{\mathbf{g}}_j^{tf}\right)^T=\mathbf{0}\ \text{ for all } i,\ j. \tag{31}$$

The proof of (28)–(29) is given in Appendix C, and the proof of (30)–(31) is given in Appendix D.

To demonstrate the performance advantage of using STFDs, we substitute (23) into (28)–(30),

$$E\left(\hat{\mathbf{s}}_i^{tf}-\mathbf{s}_i^{tf}\right)\left(\hat{\mathbf{s}}_j^{tf}-\mathbf{s}_j^{tf}\right)^H=\frac{\sigma}{N-L+1}\times$$

$$\left[\sum_{\substack{k=1\\k\neq i}}^{n_o}\frac{(\lambda_i^o-\sigma)+(\lambda_k^o-\sigma)+\frac{n_o}{L}\sigma}{(\lambda_k^o-\lambda_i^o)^2}\mathbf{s}_k^o\left(\mathbf{s}_k^o\right)^H+\sum_{k=1}^{m-n_o}\frac{(\lambda_i^o-\sigma)+\frac{n_o}{L}\sigma}{(\sigma-\lambda_i^o)^2}\mathbf{g}_k^o\left(\mathbf{g}_k^o\right)^H\right]\delta_{i,j}, \tag{32}$$

$$E\left(\hat{\mathbf{s}}_i^{tf}-\mathbf{s}_i^{tf}\right)\left(\hat{\mathbf{s}}_j^{tf}-\mathbf{s}_j^{tf}\right)^T$$

$$= -\frac{\sigma}{N-L+1}\cdot\frac{(\lambda_k^o-\sigma)+(\lambda_i^o-\sigma)+\frac{n_o}{L}\sigma}{(\lambda_k^o-\lambda_i^o)^2}\times\mathbf{s}_j^o\left(\mathbf{s}_i^o\right)^T(1-\delta_{i,j}), \tag{33}$$

and

$$E\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^{H}\hat{\mathbf{g}}_i^{tf}\right)\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^{H}\hat{\mathbf{g}}_j^{tf}\right)^{H}$$

$$= \frac{\sigma}{N-L+1}\left[\sum_{k=1}^{n_o}\frac{(\lambda_k^o-\sigma)+\frac{n_o}{L}\sigma}{(\sigma-\lambda_k^o)^2}\mathbf{s}_k^o\left(\mathbf{s}_k^o\right)^H\right]\delta_{i,j}. \tag{34}$$

From (32)–(34), two important observations are in order. First, if the signals are both localizable and separable in the time-frequency domain, then the reduction of the number of signals from n to n_o greatly reduces the estimation error, specifically when the signals are closely spaced. The examples, given in the following section, show the advantages of using t-f MUSIC with partially selected signals. The second observation relates to SNR enhancements. The above equations show that error reductions using STFDs are more pronounced for the cases of low SNR and/or closely spaced signals. It is clear from (32)–(34) that, when $\lambda_k^o \gg \sigma$ for all $k = 1, 2, ..., n_o$, the results are almost independent of L (suppose $N \gg L$ so that $N - L + 1 \simeq N$), and therefore there would be no obvious improvement in using the STFD over conventional array processing. On the other hand, when some of the eigenvalues are close to σ ($\lambda_k^o \simeq \sigma$, for some $k = 1, 2, ..., n_o$), which is the case of weak or closely spaced signals, all the results of above three equations are reduced by a factor of up to $G = L/n_o$, respectively. This factor represents, in essence, the gain achieved from using STFD processing.

IV. The Time-Frequency MUSIC

To demonstrate the robustness of the eigen-decomposition of the STFDs when used in practical applications, we consider in this section the recently proposed time-frequency MUSIC (t-f MUSIC) algorithm [7]. The DOA estimation based on time-frequency maximum likelihood (t-f ML) is investigated in reference [8].

We first recall that the DOAs are estimated in the MUSIC technique by determining the n values of θ for which the following spatial spectrum is maximized [15],

$$f_{MU}(\theta) = \left[\mathbf{a}^H(\theta)\hat{\mathbf{G}}\hat{\mathbf{G}}^H\mathbf{a}(\theta)\right]^{-1} = \left[\mathbf{a}^H(\theta)\left(\mathbf{I}-\hat{\mathbf{S}}\hat{\mathbf{S}}^H\right)\mathbf{a}(\theta)\right]^{-1}. \tag{35}$$

Where $\mathbf{a}(\theta)$ is the steering vector corresponds to θ. The variance of those estimates in the conventional MUSIC technique, assuming white noise processes, is given by [9]

$$E\left(\hat{\omega}_i-\omega_i\right)^2 = \frac{1}{2N}\frac{\mathbf{a}^H(\theta_i)\mathbf{U}\mathbf{a}(\theta_i)}{h(\theta_i)} \tag{36}$$

where ω_i is the spatial frequency associated with DOA θ_i, and $\hat{\omega}_i$ is its estimate obtained by the conventional MUSIC. In the above equation

$$\mathbf{U} = \sigma \left[\sum_{k=1}^{n} \frac{\lambda_k}{(\sigma - \lambda_k)^2} \mathbf{s}_k \mathbf{s}_k^H \right], \quad \mathbf{d}(\theta_i) = d\mathbf{a}(\theta_i)/d\omega, \quad h(\theta_i) = \mathbf{d}^H(\theta_i)\mathbf{G}\mathbf{G}^H\mathbf{d}(\theta_i). \tag{37}$$

From the results of Lemma 1, part (b), $\tilde{\mathbf{U}} = \mathbf{U}$, which implies that (36) also holds true when the conventional MUSIC algorithm is applied to FM signals in white noise.

Similarly, for t-f MUSIC with n_o signals selected, the DOAs are determined by locating the n_o peaks of the spatial spectrum defined from the n_o signals' time-frequency regions.

$$f_{MU}^{tf}(\theta) = \left[\mathbf{a}^H(\theta) \hat{\mathbf{G}}^{tf} \left(\hat{\mathbf{G}}^{tf} \right)^H \mathbf{a}(\theta) \right]^{-1} = \left[\mathbf{a}^H(\theta) \left(\mathbf{I} - \hat{\mathbf{S}}^{tf} \left(\hat{\mathbf{S}}^{tf} \right)^H \right) \mathbf{a}(\theta) \right]^{-1}. \tag{38}$$

Following the same procedure in [9] and using the results of Lemmas 2 and 3, we obtain the variance of the DOA estimates based on t-f MUSIC,

$$E\left(\hat{\omega}_i^{tf} - \omega_i \right)^2 = \frac{1}{2(N - L + 1)} \frac{\mathbf{a}^H(\theta_i)\mathbf{U}^{tf}\mathbf{a}(\theta_i)}{h^{tf}(\theta_i)} \tag{39}$$

where $\hat{\omega}_i^{tf}$ is the estimate of ω_i obtained by the t-f MUSIC, $\mathbf{U}^{tf}$ is defined in (30), and

$$h^{tf}(\theta_i) = \mathbf{d}^H(\theta_i)\mathbf{G}^{tf} \left(\mathbf{G}^{tf} \right)^H \mathbf{d}(\theta_i) \tag{40}$$

which is equal to $h(\theta_i)$, if $n_o = n$.

V. Simulation Results

Consider a uniform linear array of 8 sensors spaced by half a wavelength, and an observation period of 1024 samples. Two chirp signals emitted from two sources positioned at angle θ_1 and θ_2. The start and end frequencies of the signal source at θ_1 are $\omega_{s1} = 0$ and $\omega_{e1} = \pi$, while the corresponding two frequencies for the other source at θ_2 are $\omega_{s2} = \pi$ and $\omega_{e2} = 0$, respectively. The noise used in this simulation is zero-mean, Gaussian distributed, and temporally white. The SNR of the ith FM signal is defined as $\text{SNR}_i = 10 \log (D_i^2/\sigma)$. Fig. 3 shows the PWVD of the mixed noise-free signals for $L = 129$.

Fig. 4 displays the variance of the estimated DOA $\hat{\theta}_1$ versus SNR for the case $(\theta_1, \theta_2) = (-10°, 10°)$. The curves in this figure show the theoretical and experimental results of the conventional MUSIC and t-f MUSIC (for L=33 and 129). The Cramer-Rao bound

(CRB) is also shown in Fig. 4 for comparison. Both signals were selected when performing t-f MUSIC ($n_o = n = 2$). We assume that the number of signals is correctly estimated for each case. Simulation results were averaged over 100 independent trials of Monte Carlo experiments. The advantages of t-f MUSIC in low SNR cases are evident from this figure. The experiment results deviate from the theoretical results for low SNR, since we only considered the lowest order of the coefficients of the perturbation expansion of $\tilde{\mathbf{v}}_i$ in deriving the theoretical results (see Appendix A). Fig. 5 shows estimated spatial spectra at SNR=-20 dB based on t-f MUSIC ($L = 129$) and the conventional MUSIC. The t-f MUSIC spectral peaks are clearly resolved.

Fig. 6 shows examples of the estimated spatial spectrum based on t-f MUSIC ($L = 129$) and the conventional MUSIC where the angle separation is small ($\theta_1 = -2.5°$, $\theta_2 = 2.5°$). The input SNR is -5 dB. Two t-f MUSIC algorithms are performed using two sets of time-frequency points, each set belongs to the time-frequency signature of one source ($n_o = 1$). It is evident that the two signals cannot be resolved when MUSIC is applied, whereas by applying t-f MUSIC separately for each signal, the two signals become clearly separated and reasonable DOA estimation is achieved. This is attributed to the signal's distinct time-frequency signatures. It is noted that there is a small bias in the estimates of t-f MUSIC due to the imperfect separation of the two signals in the time-frequency domain.

It should be noted that the computation cost used to implement the t-f MUSIC is higher than the conventional MUSIC because it involves the additional processing based on bilinear t-f distributions. Nevertheless the pseudo Wigner-Ville distribution considered in this paper is relatively simple and only requires a bilinear product and one FFT operation. Moreover, there exist now several computationally efficient t-f kernels that allow time-frequency distributions to be provided via spectrogram-based implementations, recursive, and multiplication free processing. On the issue of practical implementation, many procedures have been devised so that any distribution can be calculated quickly with minimum computer resources [16]. Recently, kernels have been devised, specifically tailored to very fast computation, with the binomial kernel devised by Jeong and Williams being the prime example [17]. Methods for decomposition of kernels, leading to fast computation and increased understanding, have also been carried out by White [18], Amin

[19], [20], Venkatesan and Amin [21], and Cunningham and Williams [22].

VI. Conclusions

Subspace analyses of spatial time-frequency distribution (STFD) matrices have been presented. It has been shown that for signals with clearly defined time-frequency signatures, such as FM signals, smaller estimation errors in the signal and noise subspaces can be achieved using spatial time-frequency matrices over the subspace estimates obtained from the data covariance matrix approach. This performance improvement is the result of incorporating the time-frequency points along the instantaneous frequencies of the impinging signals on the array into the subspace estimation procedure. Under the assumption that the instantaneous frequencies are idealy located, these points belong to autoterm regions of high power concentrations, and as such, when used in constructing STFDs, they provide high SNR matrices with improved eigen-decompositions.

The advantages of STFD-based direction finding over traditional direction finding methods using data covariance matrices were demonstrated using the MUSIC algorithm. It was shown that the time-frequency MUSIC outperforms conventional MUSIC in the two situations of low SNR and closely spaced sources.

Unlike conventional array processing techniques, which are nondiscriminatory, and must therefore spatially localize all signals incident on the array, the STFD-based array processing provides the flexibility of dealing with all signal arrivals, or only a subset of them. In this respect, it does not suffer from the drawback of requiring higher number of sensors than sources. The ability to select fewer sources depends on the differences of their time-frequency signatures from those of other source signals. The eigenstructure of the STFD matrix constructed from the time-frequency points that belong to the autoterm regions of a number of sources will only yield the signal subspace of these sources. It was shown that the maximum improvement offered using STFD over data covariance matrices is obtained when constructing the STFD from only one source signal.

Appendix A

For notation simplicity, we denote $\mathbf{v}_i$, $i = 1, 2, ..., m$, as the eigenvectors of the correlation matrix $\mathbf{R}_{\mathbf{xx}}$, where the first n vectors form the signal subspace ($\mathbf{s}_i$, $i = 1, 2, ..., n$), and the

last $m-n$ vectors form the noise subspace ($\mathbf{g}_i$, $i = 1, 2, ..., m-n$).

To derive the covariance matrices, we follow the same procedure in [9] and [23], but note the fact that the underlying signals are deterministic rather than white random processes, which are considered in [9] and [23]. We define $\hat{\mathbf{R}}_{\mathbf{xx}}$ in terms of a random perturbation to $\mathbf{R}_{\mathbf{xx}}$ with a perturbation factor p, $0 < p \ll 1$. Thus,

$$\hat{\mathbf{R}}_{\mathbf{xx}} = \mathbf{R}_{\mathbf{xx}} + (\hat{\mathbf{R}}_{\mathbf{xx}} - \mathbf{R}_{\mathbf{xx}}) = \mathbf{R}_{\mathbf{xx}} + p\mathbf{B}. \tag{A.1}$$

When the source signals are FM and the noise vector forms a multivariate white Gaussian process, then $\mathbf{B}$ is a Hermitian, zero-mean random matrix whose elements are asymptotically jointly Gaussian. Let $\check{\mathbf{v}}_i$ denote the unnormalized perturbed version of the eigenvector $\mathbf{v}_i$. According to [24],

$$\check{\mathbf{v}}_i = \mathbf{v}_i + \sum_{\substack{k=1 \\ k \neq i}}^{m} \left(\sum_{l=1}^{\infty} t_{lk}^{(i)} p^l \right) \mathbf{v}_k \tag{A.2}$$

where $t_{lk}^{(i)}$, $l = 1, 2, ...,$ are the coefficients of the perturbation expansion of $\check{\mathbf{v}}_i$ along $\mathbf{v}_k$. By keeping the term with the lowest order of p, then [23]

$$t_{1k}^{(i)} = \frac{\mathbf{v}_k^H \mathbf{B} \mathbf{v}_i}{\lambda_k - \lambda_i}, \; k \neq i. \tag{A.3}$$

The mean square value of $t_{1k}^{(i)}$ is given by

$$E\left[\left| t_{1k}^{(i)} \right|^2 \right] = E\left[\frac{\mathbf{v}_k^H \mathbf{B} \mathbf{v}_i \mathbf{v}_i^H \mathbf{B} \mathbf{v}_k}{(\lambda_k - \lambda_i)^2} \right]. \tag{A.4}$$

To evaluate the numerator in the above equation, we consider the following general case.

$$\begin{aligned} E\left[\mathbf{v}_i^H \mathbf{B} \mathbf{v}_j \mathbf{v}_k^H \mathbf{B} \mathbf{v}_l \right] &= \frac{1}{p^2} E\left[\mathbf{v}_i^H \left(\hat{\mathbf{R}}_{\mathbf{xx}} - \mathbf{R}_{\mathbf{xx}} \right) \mathbf{v}_j \mathbf{v}_k^H \left(\hat{\mathbf{R}}_{\mathbf{xx}} - \mathbf{R}_{\mathbf{xx}} \right) \mathbf{v}_l \right] \\ &= \frac{1}{(Np)^2} E\left[\left(\sum_{r=1}^{N} \mathbf{v}_i^H \mathbf{x}(t_r) \mathbf{x}^H(t_r) \mathbf{v}_j \right) \left(\sum_{q=1}^{N} \mathbf{v}_k^H \mathbf{x}(t_q) \mathbf{x}^H(t_q) \mathbf{v}_l \right) \right] \\ &\quad - \frac{1}{p^2} \mathbf{v}_i^H \mathbf{R}_{\mathbf{xx}} \mathbf{v}_j \mathbf{v}_k^H \mathbf{R}_{\mathbf{xx}} \mathbf{v}_l. \end{aligned} \tag{A.5}$$

It can be easily realized that the expected value in (A.5) is taken from a product of four non-zero mean Gaussian random variables. It is well known that for Gaussian random variable x_1, x_2, x_3, x_4 with non-zero means,

$$\begin{aligned} E[x_1 x_2 x_3 x_4] &= E[x_1 x_2] E[x_3 x_4] + E[x_1 x_3] E[x_2 x_4] \\ &\quad + E[x_1 x_4] E[x_2 x_3] - 2E[x_1] E[x_2] E[x_3] E[x_4]. \end{aligned} \tag{A.6}$$

Using the properties of the zero-mean circular complex Gaussian noise vector and the deterministic source signal vector, then

$$\begin{aligned} E[\mathbf{x}(t_r)] &= \mathbf{y}(t_r), \\ E\left[\mathbf{x}(t_r)\mathbf{x}^H(t_q)\right] &= \mathbf{y}(t_r)\mathbf{y}^H(t_q) + \sigma\mathbf{I}\delta_{r,q}, \\ E\left[\mathbf{x}(t_r)\mathbf{x}^T(t_q)\right] &= \mathbf{y}(t_r)\mathbf{y}^T(t_q). \end{aligned}$$

Accordingly, (A.5) can be written as

$$\begin{aligned} & E\left[\mathbf{v}_i^H\mathbf{B}\mathbf{v}_j\mathbf{v}_k^H\mathbf{B}\mathbf{v}_l\right] \\ &= \frac{1}{(Np)^2}\sum_{r=1}^{N}\sum_{q=1}^{N} E\left[\mathbf{v}_i^H\mathbf{x}(t_r)\mathbf{x}^H(t_r)\mathbf{v}_j\right] E\left[\mathbf{v}_k^H\mathbf{x}(t_q)\mathbf{x}^H(t_q)\mathbf{v}_l\right] \\ &\quad + \frac{1}{(Np)^2}\sum_{r=1}^{N}\sum_{q=1}^{N} E\left[\mathbf{v}_i^H\mathbf{x}(t_r)\mathbf{v}_k^H\mathbf{x}(t_q)\right] E\left[\mathbf{x}^H(t_r)\mathbf{v}_j\mathbf{x}^H(t_q)\mathbf{v}_l\right] \\ &\quad + \frac{1}{(Np)^2}\sum_{r=1}^{N}\sum_{q=1}^{N} E\left[\mathbf{v}_i^H\mathbf{x}(t_r)\mathbf{x}^H(t_q)\mathbf{v}_l\right] E\left[\mathbf{v}_k^H\mathbf{x}(t_q)\mathbf{x}^H(t_r)\mathbf{v}_j\right] \\ &\quad - 2\frac{1}{(Np)^2}\sum_{r=1}^{N}\sum_{q=1}^{N} E\left[\mathbf{v}_i^H\mathbf{x}(t_r)\right] E\left[\mathbf{x}^H(t_r)\mathbf{v}_j\right] E\left[\mathbf{v}_k^H\mathbf{x}(t_q)\right] E\left[\mathbf{x}^H(t_q)\mathbf{v}_l\right] \\ &\quad - \frac{1}{p^2}\mathbf{v}_i^H\mathbf{R}_{\mathbf{xx}}\mathbf{v}_j\mathbf{v}_k^H\mathbf{R}_{\mathbf{xx}}\mathbf{v}_l \\ &= \frac{1}{(Np)^2}\sum_{r=1}^{N}\sum_{q=1}^{N}\left[\mathbf{v}_i^H\mathbf{v}_l\sigma\delta_{r,q}\mathbf{v}_k^H\mathbf{y}(t_q)\mathbf{y}^H(t_r)\mathbf{v}_j\right] \\ &\quad + \frac{1}{(Np)^2}\sum_{r=1}^{N}\sum_{q=1}^{N}\left[\mathbf{v}_i^H\mathbf{y}(t_r)\mathbf{y}^H(t_q)\mathbf{v}_l\mathbf{v}_k^H\mathbf{v}_j\sigma\delta_{r,q}\right] \\ &\quad + \frac{1}{(Np)^2}\sum_{r=1}^{N}\sum_{q=1}^{N}\left[\mathbf{v}_i^H\mathbf{v}_l\mathbf{v}_k^H\mathbf{v}_j\sigma^2\delta_{r,q}\right] \\ &= \frac{1}{(Np)^2}\sum_{r=1}^{N}\left[\delta_{i,l}\sigma\mathbf{v}_k^H\mathbf{y}(t_r)\mathbf{y}^H(t_r)\mathbf{v}_j\right] \\ &\quad + \frac{1}{(Np)^2}\sum_{r=1}^{N}\left[\mathbf{v}_i^H\mathbf{y}(t_r)\mathbf{y}^H(t_r)\mathbf{v}_l\delta_{j,k}\sigma\right] + \frac{1}{(Np)^2}\sum_{r=1}^{N}\left[\delta_{i,l}\delta_{j,k}\sigma^2\right]. \end{aligned} \tag{A.7}$$

By using the uncorrelation assumption (7),

$$\begin{aligned} \frac{1}{N}\sum_{r=1}^{N}\mathbf{y}(t_r)\mathbf{y}^H(t_r) &= \mathbf{A}\left[\frac{1}{N}\sum_{r=1}^{N}\mathbf{d}(t_r)\mathbf{d}^H(t_r)\right]\mathbf{A}^H \\ &= \mathbf{A}\mathbf{R}_{\mathbf{dd}}\mathbf{A}^H = \mathbf{R}_{\mathbf{yy}} = \mathbf{R}_{\mathbf{xx}} - \sigma\mathbf{I}, \end{aligned} \tag{A.8}$$

equation (A.7) simplifies to

$$
\begin{aligned}
& E\left[\mathbf{v}_i^H \mathbf{B} \mathbf{v}_j \mathbf{v}_k^H \mathbf{B} \mathbf{v}_l\right] \\
&= \frac{\sigma}{Np^2}\left[\delta_{i,l}\mathbf{v}_k^H \mathbf{R}_{\mathbf{yy}} \mathbf{v}_j + \mathbf{v}_i^H \mathbf{R}_{\mathbf{yy}} \mathbf{v}_l \delta_{j,k}\right] + \frac{\sigma^2}{Np^2}\delta_{i,l}\delta_{j,k} \\
&= \frac{\sigma}{Np^2}\left[\delta_{i,l}\mathbf{v}_k^H \left(\mathbf{R}_{\mathbf{xx}} - \sigma\mathbf{I}\right)\mathbf{v}_j + \mathbf{v}_i^H \left(\mathbf{R}_{\mathbf{xx}} - \sigma\mathbf{I}\right)\mathbf{v}_l \delta_{j,k}\right] + \frac{\sigma^2}{Np^2}\delta_{i,l}\delta_{j,k} \\
&= \frac{\sigma}{Np^2}\left[\delta_{i,l}\mathbf{v}_k^H \mathbf{R}_{\mathbf{xx}} \mathbf{v}_j + \mathbf{v}_i^H \mathbf{R}_{\mathbf{xx}} \mathbf{v}_l \delta_{j,k} - \sigma\delta_{i,l}\delta_{j,k}\right] \\
&= \frac{\sigma}{Np^2}\left[\lambda_i + \lambda_j - \sigma\right]\delta_{i,l}\delta_{j,k}. \qquad \text{(A.9)}
\end{aligned}
$$

Therefore,

$$
E\left[\left|t_{1k}^{(i)}\right|^2\right] = E\left[\frac{\mathbf{v}_k^H \mathbf{B} \mathbf{v}_i \mathbf{v}_i^H \mathbf{B} \mathbf{v}_k}{(\lambda_k - \lambda_i)^2}\right] = \frac{\sigma}{Np^2}\frac{(\lambda_i + \lambda_k - \sigma)}{(\lambda_i - \lambda_k)^2},\ k \neq i, \qquad \text{(A.10)}
$$

and

$$
E\left[t_{1k}^{(i)}(t_{1k}^{(j)})^*\right] = E\left[\frac{\mathbf{v}_k^H \mathbf{B} \mathbf{v}_i \mathbf{v}_j^H \mathbf{B} \mathbf{v}_k}{(\lambda_k - \lambda_i)(\lambda_k - \lambda_j)}\right] = 0,\ k \neq i,\ k \neq j. \qquad \text{(A.11)}
$$

It is shown in [13] that

$$
cov\left(\hat{\mathbf{v}}_i, \hat{\mathbf{v}}_j\right) = cov\left(\check{\mathbf{v}}_i, \check{\mathbf{v}}_j\right) + o(N^{-2}).
$$

By ignoring the terms of N^{-2}, then

$$
\begin{aligned}
cov\left(\hat{\mathbf{v}}_i, \hat{\mathbf{v}}_j\right) &\simeq cov\left(\check{\mathbf{v}}_i, \check{\mathbf{v}}_j\right) \simeq E\left[\left(\sum_{\substack{k=1\\k\neq i}}^{m} t_{1k}^{(i)} p \mathbf{v}_k\right)\left(\sum_{\substack{k=1\\k\neq j}}^{m} t_{1k}^{(j)} p \mathbf{v}_k\right)^H\right] \\
&= E\left[\sum_{\substack{k=1\\k\neq i}}^{m}\left|t_{1k}^{(i)}\right|^2 p^2 \mathbf{v}_k \mathbf{v}_k^H \delta_{i,j}\right] = \frac{\sigma}{N}\sum_{\substack{k=1\\k\neq i}}^{m}\frac{\lambda_i + \lambda_k - \sigma}{(\lambda_i - \lambda_k)^2}\mathbf{v}_k \mathbf{v}_k^H \delta_{i,j}. \qquad \text{(A.12)}
\end{aligned}
$$

Replacing $\mathbf{v}_k$ by $\mathbf{s}_k$ or $\mathbf{g}_k$ leads to equation (8). Similarly,

$$
\begin{aligned}
cov\left(\hat{\mathbf{v}}_i, \hat{\mathbf{v}}_j^*\right) &\simeq cov\left(\check{\mathbf{v}}_i, \check{\mathbf{v}}_j^*\right) \\
&\simeq E\left[\left(\sum_{\substack{k=1\\k\neq i}}^{m} t_{1k}^{(i)} p \mathbf{v}_k\right)\left(\sum_{\substack{k=1\\k\neq j}}^{m} t_{1k}^{(j)} p \mathbf{v}_k\right)^T\right] = p^2 E\left[\sum_{\substack{k_1=1\\k_1\neq i}}^{m}\sum_{\substack{k_2=1\\k_2\neq j}}^{m} t_{1k_1}^{(i)} t_{1k_2}^{(j)} \mathbf{v}_{k_1} \mathbf{v}_{k_2}^T\right] \\
&= p^2 \sum_{\substack{k_1=1\\k_1\neq i}}^{m}\sum_{\substack{k_2=1\\k_2\neq j}}^{m} E\left[\frac{\mathbf{v}_{k_1}^H \mathbf{B} \mathbf{v}_i}{(\lambda_{k_1} - \lambda_i)}\frac{\mathbf{v}_{k_2}^H \mathbf{B} \mathbf{v}_j}{(\lambda_{k_2} - \lambda_j)}\mathbf{v}_{k_1} \mathbf{v}_{k_2}^T\right]. \qquad \text{(A.13)}
\end{aligned}
$$

From (A.9), it is clear that the above equation has non-zero value only when $k_1 = j$ and $k_2 = i$. Noting the fact that $k_1 \neq i$ and $k_2 \neq j$, (A.13) becomes

$$\begin{aligned} cov\left(\hat{\mathbf{v}}_i, \hat{\mathbf{v}}_j^*\right) &\simeq -\frac{p^2}{(\lambda_j - \lambda_i)^2} E\left[\mathbf{v}_j^H \mathbf{B}\mathbf{v}_i \mathbf{v}_i^H \mathbf{B}\mathbf{v}_j\right] \mathbf{v}_j \mathbf{v}_i^T (1 - \delta_{i,j}) \\ &= -\frac{\sigma}{N} \frac{\lambda_i + \lambda_j - \sigma}{(\lambda_j - \lambda_i)^2} \mathbf{v}_j \mathbf{v}_i^T (1 - \delta_{i,j}). \end{aligned} \tag{A.14}$$

For the signal subspace, $\mathbf{v}_i$ is $\mathbf{s}_i$, $i = 1, ..., n$, and (A.14) yields (9) and this concludes the proof for part a) of Lemma 1.

Appendix B

Using eigendecomposition theory, we have

$$\mathbf{A}^o \mathbf{R}_{\mathbf{dd}}^o (\mathbf{A}^o)^H = \mathbf{Q}\Xi\mathbf{Q}^H \tag{B.1}$$

where $\Xi = diag\left[\xi_1, \cdots, \xi_{n_o}, 0, \cdots, 0\right]$ is a diagonal matrix whose elements are the eigenvalues of $\mathbf{A}^o \mathbf{R}_{\mathbf{dd}}^o (\mathbf{A}^o)^H$ and $\mathbf{Q}$ is the corresponding eigenvector matrix. Clearly, $\xi_i = \lambda_i^o - \sigma, i = 1, \cdots, n_o$, and $\mathbf{Q} = [\ \mathbf{S}^o \ | \ \mathbf{G}^o\ ]$.

From the definition of λ_i^o and λ_i^{tf}, it is evident that

$$\mathbf{D} = \frac{L}{n_o}\mathbf{Q}\Xi\mathbf{Q}^H + \sigma\mathbf{I} = \mathbf{Q}\left[\frac{L}{n_o}\Xi + \sigma\mathbf{I}\right]\mathbf{Q}^H \tag{B.2}$$

Therefore, $\mathbf{R}_{\mathbf{xx}}^o$ and $\mathbf{D}$ share the same set of eigenvectors, that proves part a) of Lemma 1. The ith eigenvalue of $\mathbf{D}$ is $\frac{L}{n_o}\xi_i + \sigma = \frac{L}{n_o}(\lambda_i^o - \sigma) + \sigma$ for $i \leq n_o$ and is σ for $n_o < i \leq m$, subsequently leading to part b).

Appendix C

Similar to Appendix A, we let $\mathbf{v}_i$, $i = 1, 2, ..., m$, represent the whole eigenvectors of the STFD matrix $\mathbf{D}$, where the first n_o vectors form the signal subspace ($\mathbf{s}_i^{tf}$, $i = 1, 2, ..., n_o$), while the last $m - n_o$ vectors form the noise subspace ($\mathbf{g}_i^{tf}$, $i = 1, 2, ..., m - n_o$). As discussed in Section III, we assume that the selected time-frequency points belong to regions where no crossterm components are present.

For an array mixture of FM signals, we select points from n_o signals at the time-frequency domain, where the pseudo Wigner-Ville distribution matrix is defined in (13).

We define $\hat{\mathbf{D}}$ in terms of a random perturbation to $\mathbf{D}$ with a perturbation factor p, $0 < p \ll 1$. Thus,

$$\hat{\mathbf{D}} = \mathbf{D} + (\hat{\mathbf{D}} - \mathbf{D}) = \mathbf{D} + p\mathbf{B}. \tag{C.1}$$

Matrix $\mathbf{B}$ is a Hermitian, zero-mean random matrix whose elements are asymptotically Jointly Gaussian [8]. Similar to Appendix A, we derive

$$\begin{aligned}
& E\left[\mathbf{v}_i^H \mathbf{B}\mathbf{v}_j \mathbf{v}_k^H \mathbf{B}\mathbf{v}_l\right] \\
&= \frac{1}{p^2} E\left[\mathbf{v}_i^H \left(\hat{\mathbf{D}} - \mathbf{D}\right)\mathbf{v}_j \mathbf{v}_k^H \left(\hat{\mathbf{D}} - \mathbf{D}\right)\mathbf{v}_l\right] \\
&= \frac{1}{\left(n_o p(N-L+1)\right)^2} E\left[\left(\sum_{q=1}^{n_o}\sum_{i=1}^{N-L+1} \mathbf{v}_i^H \mathbf{D}_{\mathbf{xx}}(t_i, f_{q,i})\mathbf{v}_j\right)\left(\sum_{q=1}^{n_o}\sum_{i=1}^{N-L+1} \mathbf{v}_k^H \mathbf{D}_{\mathbf{xx}}(t_i, f_{q,i})\mathbf{v}_l\right)\right] \\
&\quad -\frac{1}{p^2}\mathbf{v}_i^H \mathbf{D}\mathbf{v}_j \mathbf{v}_k^H \mathbf{D}\mathbf{v}_l.
\end{aligned} \tag{C.2}$$

Substituting (21) and (A.7) into (C.2), we obtain

$$\begin{aligned}
& E\left[\mathbf{v}_i^H \mathbf{B}\mathbf{v}_j \mathbf{v}_k^H \mathbf{B}\mathbf{v}_l\right] \\
&= \frac{1}{\left(n_o p(N-L+1)\right)^2} \sum_{q_1=1}^{n_o}\sum_{q_2=1}^{n_o}\sum_{i_1=1}^{N-L+1}\sum_{i_2=1}^{N-L+1}\sum_{\tau_1=-\frac{L-1}{2}}^{\frac{L-1}{2}}\sum_{\tau_2=-\frac{L-1}{2}}^{\frac{L-1}{2}} e^{-j4\pi\left[f_{q_1,i_1}\tau_1 + f_{q_2,i_2}\tau_2\right]} \\
&\quad \times\Big\{E\left[\mathbf{v}_i^H \mathbf{x}(t_{i_1}+\tau_1)\mathbf{x}^H(t_{i_1}-\tau_1)\mathbf{v}_j\right] E\left[\mathbf{v}_k^H \mathbf{x}(t_{i_2}+\tau_2)\mathbf{x}^H(t_{i_2}-\tau_2)\mathbf{v}_l\right] \\
&\quad +E\left[\mathbf{v}_i^H \mathbf{x}(t_{i_1}+\tau_1)\mathbf{v}_k^H \mathbf{x}(t_{i_2}+\tau_2)\right] E\left[\mathbf{x}^H(t_{i_1}-\tau_1)\mathbf{v}_j \mathbf{x}^H(t_{i_2}-\tau_2)\mathbf{v}_l\right] \\
&\quad +E\left[\mathbf{v}_i^H \mathbf{x}(t_{i_1}+\tau_1)\mathbf{x}^H(t_{i_2}-\tau_2)\mathbf{v}_l\right] E\left[\mathbf{v}_k^H \mathbf{x}(t_{i_2}+\tau_2)\mathbf{x}^H(t_{i_1}-\tau_1)\mathbf{v}_j\right] \\
&\quad -2E\left[\mathbf{v}_i^H \mathbf{x}(t_{i_1}+\tau_1)\right] E\left[\mathbf{x}^H(t_{i_1}-\tau_1)\mathbf{v}_j\right] E\left[\mathbf{v}_k^H \mathbf{x}(t_{i_2}+\tau_2)\right] E\left[\mathbf{x}^H(t_{i_2}-\tau_2)\mathbf{v}_l\right]\Big\} \\
&\quad -\frac{1}{p^2}\mathbf{v}_i^H \mathbf{D}\mathbf{v}_j \mathbf{v}_k^H \mathbf{v}_l \\
&= \frac{1}{\left(n_o p(N-L+1)\right)^2} \sum_{q_1=1}^{n_o}\sum_{q_2=1}^{n_o}\sum_{i_1=1}^{N-L+1}\sum_{i_2=1}^{N-L+1}\sum_{\tau_1=-\frac{L-1}{2}}^{\frac{L-1}{2}}\sum_{\tau_2=-\frac{L-1}{2}}^{\frac{L-1}{2}} e^{-j4\pi\left[f_{q_1,i_1}\tau_1 + f_{q_2,i_2}\tau_2\right]} \\
&\quad \times\Big[\mathbf{v}_i^H \mathbf{y}(t_{i_1}+\tau_1)\mathbf{y}^H(t_{i_2}-\tau_2)\mathbf{v}_l \sigma\delta_{j,k}\delta_{t_{i_1}-\tau_1,t_{i_2}+\tau_2} \\
&\quad +\sigma\delta_{i,l}\delta_{t_{i_1}+\tau_1,t_{i_2}-\tau_2}\mathbf{v}_k^H \mathbf{y}(t_{i_2}+\tau_2)\mathbf{y}^H(t_{i_1}-\tau_1)\mathbf{v}_j \\
&\quad +\sigma^2\delta_{i,l}\delta_{j,k}\delta_{t_{i_1},t_{i_2}}\delta_{\tau_1,\tau_2}\Big].
\end{aligned} \tag{C.3}$$

Under the assumption of no crossterms, q_1 should be equivalent to q_2 to have non-zero values, and in this case, $q_1 = q_2 = q$. Note that within the time-frequency region of the qth signal, $\mathbf{y}(t) = \mathbf{y}_q(t) \underline{\underline{\triangle}} \mathbf{A}\mathbf{d}_q(t)$. When the third-order derivative of the phase is negligible over $[t-L+1, t+L-1]$ for any signal and any t, we have

$$
\begin{aligned}
& E\left[\mathbf{v}_i^H \mathbf{B}\mathbf{v}_j \mathbf{v}_k^H \mathbf{B}\mathbf{v}_l\right] \\
&\simeq \frac{1}{(Np(N-L+1))^2} \sum_{q=1}^{n_o} \sum_{i_1=1}^{N-L+1} \sum_{i_2=1}^{N-L+1} \sum_{\tau_1=-\frac{L-1}{2}}^{\frac{L-1}{2}} \sum_{\tau_2=-\frac{L-1}{2}}^{\frac{L-1}{2}} \\
&\quad \times \Big[\mathbf{v}_i^H \mathbf{y}(t_{i_1}+\tau_1)\mathbf{y}^H(t_{i_2}-\tau_2)\mathbf{v}_l \sigma \delta_{j,k} \delta_{t_{i_1}-\tau_1, t_{i_2}+\tau_2} \\
&\quad + \sigma \delta_{i,l} \delta_{t_{i_1}+\tau_1, t_{i_1}-\tau_1} \mathbf{v}_k^H \mathbf{y}(t_{i_2}+\tau_2)\mathbf{y}^H(t_{i_1}-\tau_1)\mathbf{v}_j\Big] \\
&\quad + \frac{\sigma^2 L}{n_o(N-L+1)p^2} \delta_{i,l}\delta_{j,k} \\
&\simeq \frac{\sigma L}{n_o(N-L+1)p^2}\left[\frac{L}{n_o}(\lambda_i - \sigma) + (\lambda_j - \sigma) + \sigma\right] \delta_{i,l}\delta_{j,k} \\
&= \frac{\sigma L}{n_o(N-L+1)p^2}\left[(\lambda_i^{tf} + \lambda_j^{tf} - \sigma)\right] \delta_{i,l}\delta_{j,k}. \qquad \text{(C.4)}
\end{aligned}
$$

Let $\tilde{\mathbf{v}}_i$ denote the unnormalized eigenvector given in a perturbation expansions by

$$
\tilde{\mathbf{v}}_i = \mathbf{v}_i + \sum_{\substack{k=1 \\ k\neq i}}^{m} \left(\sum_{l=1}^{\infty} t_{lk}^{(i)} p^l\right) \mathbf{v}_k \qquad \text{(C.5)}
$$

where $t_{lk}^{(i)}$, $l = 1, 2, ...,$ are the coefficients of the perturbation expansion of $\tilde{\mathbf{v}}_i$ along $\mathbf{v}_k$, and keeping the term with the lowest order of p, then

$$
t_{1k}^{(i)} = \frac{\mathbf{v}_k^H \mathbf{B}\mathbf{v}_i}{\lambda_k^{tf} - \lambda_i^{tf}}, \; k \neq i. \qquad \text{(C.6)}
$$

Therefore,

$$
\begin{aligned}
E\left[\left|t_{1k}^{(i)}\right|^2\right] &= E\left[\frac{\mathbf{v}_k^H \mathbf{B}\mathbf{v}_i \mathbf{v}_i^H \mathbf{B}\mathbf{v}_k}{(\lambda_k^{tf} - \lambda_i^{tf})^2}\right] \\
&= \frac{\sigma L}{n_o(N-L+1)p^2} \frac{(\lambda_i^{tf} + \lambda_k^{tf} - \sigma)}{\left(\lambda_i^{tf} - \lambda_k^{tf}\right)^2}, \; k \neq i, \qquad \text{(C.7)}
\end{aligned}
$$

and

$$
E\left[t_{1k}^{(i)}(t_{1k}^{(j)})^*\right] = E\left[\frac{\mathbf{v}_k^H \mathbf{B}\mathbf{v}_i \mathbf{v}_j^H \mathbf{B}\mathbf{v}_k}{(\lambda_k^{tf} - \lambda_i^{tf})(\lambda_k^{tf} - \lambda_j^{tf})}\right] = 0, \; k \neq i, \; k \neq j. \qquad \text{(C.8)}
$$

Similar to Appendix A, we follow

$$\begin{aligned} cov\left(\hat{\mathbf{v}}_i, \hat{\mathbf{v}}_j\right) &\simeq cov\left(\check{\mathbf{v}}_i, \check{\mathbf{v}}_j\right) \\ &\simeq E\left[\left(\sum_{\substack{k=1\\k\neq i}}^{m} t_{1k}^{(i)} p\mathbf{v}_k\right)\left(\sum_{\substack{k=1\\k\neq j}}^{m} t_{1k}^{(j)} p\mathbf{v}_k\right)^H\right] = E\left[\sum_{\substack{k=1\\k\neq i}}^{m} \left|t_{1k}^{(i)}\right|^2 p^2 \mathbf{v}_k\mathbf{v}_k^H \delta_{i,j}\right] \\ &= \frac{\sigma L}{n_o(N-L+1)} \sum_{\substack{k=1\\k\neq i}}^{m} \frac{\lambda_i^{tf} + \lambda_k^{tf} - \sigma}{(\lambda_i^{tf} - \lambda_k^{tf})^2} \mathbf{v}_k\mathbf{v}_k^H \delta_{i,j}. \end{aligned} \tag{C.9}$$

(28) follows by properly replacing $\mathbf{v}_k$ by $\mathbf{s}_k^{tf}$ or $\mathbf{g}_k^{tf}$. Similarly,

$$\begin{aligned} cov\left(\hat{\mathbf{v}}_i, \hat{\mathbf{v}}_j^*\right) &\simeq cov\left(\check{\mathbf{v}}_i, \check{\mathbf{v}}_j^*\right) \simeq E\left[\left(\sum_{\substack{k=1\\k\neq i}}^{m} t_{1k}^{(i)} p\mathbf{v}_k\right)\left(\sum_{\substack{k=1\\k\neq j}}^{m} t_{1k}^{(j)} p\mathbf{v}_k\right)^T\right] \\ &= p^2 \sum_{\substack{k_1=1\\k_1\neq i}}^{m} \sum_{\substack{k_2=1\\k_2\neq j}}^{m} E\left[\frac{\mathbf{v}_{k_1}^H \mathbf{B}\mathbf{v}_i}{(\lambda_{k_1}^{tf} - \lambda_i^{tf})} \frac{\mathbf{v}_{k_2}^H \mathbf{B}\mathbf{v}_j}{(\lambda_{k_2}^{tf} - \lambda_j^{tf})} \mathbf{v}_{k_1}\mathbf{v}_{k_2}^T\right] \\ &= -\frac{p^2}{(\lambda_j^{tf} - \lambda_i^{tf})^2} E\left[\mathbf{v}_j^H \mathbf{B}\mathbf{v}_i \mathbf{v}_i^H \mathbf{B}\mathbf{v}_j\right] \mathbf{v}_j\mathbf{v}_i^T (1-\delta_{i,j}) \\ &= -\frac{\sigma L}{n_o(N-L+1)} \frac{\lambda_i^{tf} + \lambda_j^{tf} - \sigma}{(\lambda_j^{tf} - \lambda_i^{tf})^2} \mathbf{v}_j\mathbf{v}_i^T (1-\delta_{i,j}). \end{aligned} \tag{C.10}$$

For the columns of signal subspace, $\mathbf{v}_i$ becomes $\mathbf{s}_i^{tf}$, and (C.10) becomes (29).

Appendix D

This appendix follows the procedure of [9]. Denote

$$\mathbf{\Gamma} = \left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{D}}\mathbf{G}^{tf},$$

and γ_i the ith column of $\mathbf{\Gamma}$. Using the results of (C.2) – (C.4), and the fact $\left(\mathbf{S}^{tf}\right)^H \mathbf{D}\mathbf{G}^{tf} = \mathbf{0}$, we have

$$E\left[\gamma_i\gamma_j^H\right]_{t,q} = E\left[\left(\left(\mathbf{s}_t^{tf}\right)^H \hat{\mathbf{D}}\mathbf{g}_i^{tf}\right)\left(\left(\mathbf{g}_j^{tf}\right)^H \hat{\mathbf{D}}\mathbf{s}_q^{tf}\right)\right] = \frac{\sigma L}{n_o(N-L+1)} \lambda_t^{tf} \delta_{t,q}\delta_{i,j}. \tag{D.1}$$

Subsequently

$$E\left[\gamma_i\gamma_j^H\right] = E\left[\left(\left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{D}}\mathbf{g}_i^{tf}\right)\left(\left(\mathbf{g}_j^{tf}\right)^H \hat{\mathbf{D}}\mathbf{S}^{tf}\right)\right] = \frac{\sigma L}{n_o(N-L+1)} \mathbf{\Lambda}^{tf} \delta_{i,j}, \tag{D.2}$$

where $\mathbf{\Lambda}^{tf} = diag\left[\lambda_1^{tf}, ..., \lambda_{n_o}^{tf}\right]$. Similarly,

$$E\left[\gamma_i \gamma_j^T\right]_{t,q} = E\left[\left(\left(\mathbf{s}_t^{tf}\right)^H \hat{\mathbf{D}}\mathbf{g}_i^{tf}\right)\left(\left(\mathbf{s}_q^{tf}\right)^H \hat{\mathbf{D}}\mathbf{g}_j^{tf}\right)\right] = 0, \tag{D.3}$$

and subsequently

$$E\left[\gamma_i \gamma_j^T\right] = \mathbf{0}. \tag{D.4}$$

Since $\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{g}}_i^{tf}$ has the same limiting distribution as that of $-\mathbf{S}^{tf}(\mathbf{\Gamma} - \sigma\mathbf{I})^{-1}\gamma_i$ [8], then it follows

$$\begin{aligned} & E\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{g}}_i^{tf}\right)\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{g}}_j^{tf}\right)^H \\ &= \frac{\sigma L}{n_o(N-L+1)}\left[\mathbf{S}(\mathbf{\Lambda}^{tf} - \sigma\mathbf{I})^{-1}\mathbf{\Lambda}^{tf}(\mathbf{\Lambda}^{tf} - \sigma\mathbf{I})^{-1}\mathbf{S}^H\right]\delta_{i,j} \\ &= \frac{\sigma L}{n_o(N-L+1)}\left[\sum_{k=1}^{n_o} \frac{\lambda_k^{tf}}{(\sigma - \lambda_k^{tf})^2}\mathbf{s}_k^{tf}\left(\mathbf{s}_k^{tf}\right)^H\right]\delta_{i,j}, \end{aligned} \tag{D.5}$$

and

$$E\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{g}}_i^{tf}\right)\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{g}}_j^{tf}\right)^T = \mathbf{0} \text{ for all } i,\ j. \tag{D.6}$$

Acknowledgment

The authors would like to thank Dr. Kehu Yang of ATR Adaptive Communications Research Laboratories, Japan, for his helpful discussions on this subject during his visit to Villanova University.

References

[1] L. Cohen, *Time-Frequency Analysis*, Englewood Cliffs, NJ: Prentice Hall, 1995.

[2] S. Qian and D. Chen, *Joint Time-Frequency Analysis*, Englewood Cliffs, NJ: Prentice Hall, 1996.

[3] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing*, S. Haykin ed., vol. I, Englewood Cliffs, NJ: Prentice-Hall, 1990.

[4] L. Cohen, "Time-frequency signal analysis – a review," *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, July 1989.

[5] A. Belouchrani and M. Amin, "Blind source separation based on time-frequency signal representation," *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. 2888–2898, Nov. 1998.

[6] M. G. Amin, "Spatial time-frequency distributions for direction finding and blind source separation," (invited paper) in *Proc. SPIE: Wavelet Applications IV*, vol. 3723, pp. 62–70, April 1999.

[7] A. Belouchrani and M. Amin, "Time-frequency MUSIC," *IEEE Signal Processing Letters*, vol. 6, no. 5, pp. 109–110, May 1999.

[8] Y. Zhang, W. Mu, and M. G. Amin, "Time-frequency maximum likelihood methods for direction finding," *Journal of Franklin Institute*, vol. 337, no. 4, pp. 483–497, July 2000.

[9] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, no. 5, pp. 720–741, May 1989.

[10] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd edition. Baltimore, MD: Johns Hopkins University Press, 1996.

[11] K. Sekihara, S. Nagarajan, D. Poeppel, and Y. Miyashita, "Time-frequency MEG-MUSIC algorithm," *IEEE Trans. Medical Imaging*, vol. 18, no. 1, pp. 92–97, Jan. 1999.

[12] J. E. Hudson, *Adaptive Array Principles*. London: Peter Peregrinus Ltd., 1981.

[13] L. C. Godara, "Application of antenna arrays to mobile communications, part II: Beam-forming and direction-of-arrival considerations," *Proc. IEEE*, vol. 85, no. 8, pp. 1195–1245, Aug. 1997.

[14] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE*

Trans. Acoust., Speech, Signal Proc., vol. 33, no. 2, pp. 387–392, April 1985.

[15] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, March 1986.

[16] L. Cohen, "Recent developments in the core of digital signal processing: time-frequency analysis," *IEEE Signal Processing Magazine*, vol. 16, no. 1, pp. 22–28, Jan. 1999.

[17] J. Jeong and W. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Processing*, vol. 40, no. 2, pp. 402–412, Feb. 1992.

[18] L. White, "Transition kernels for bilinear time-frequency signal representations," *IEEE Trans. Signal Processing*, vol. 39, no. 2, pp. 542–544, Feb. 1991.

[19] M. Amin, "Spectral decomposition of the time-frequency distribution kernels," *IEEE Trans. Signal Processing*, vol. 42, no. 5, pp. 1156–1166, May 1994.

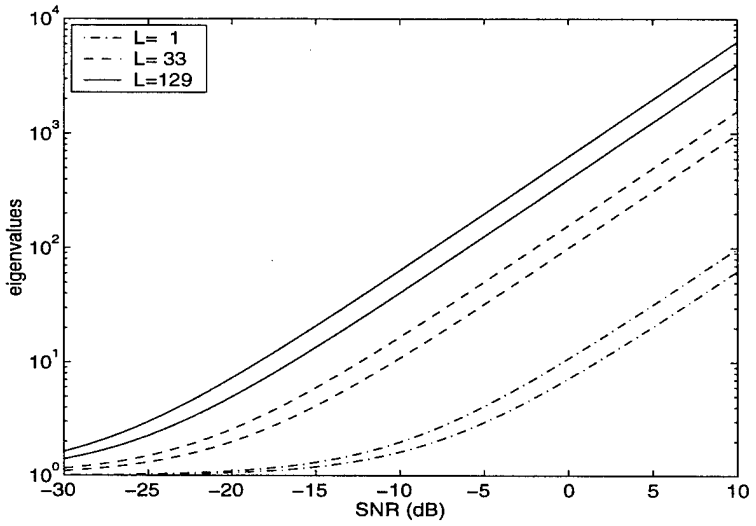
[20] M. Amin, "Recursive kernels for time-frequency signal representations," *IEEE Trans. Signal Processing*, vol. 3, no. 1, pp. 16–18, Jan. 1996.

[21] G. Venkatesan and M. Amin, "Time-frequency distribution kernel design over a discrete powers-of-two space," *IEEE Signal Processing Letters*, vol. 3, no. 12, pp. 305–306, Dec. 1996.

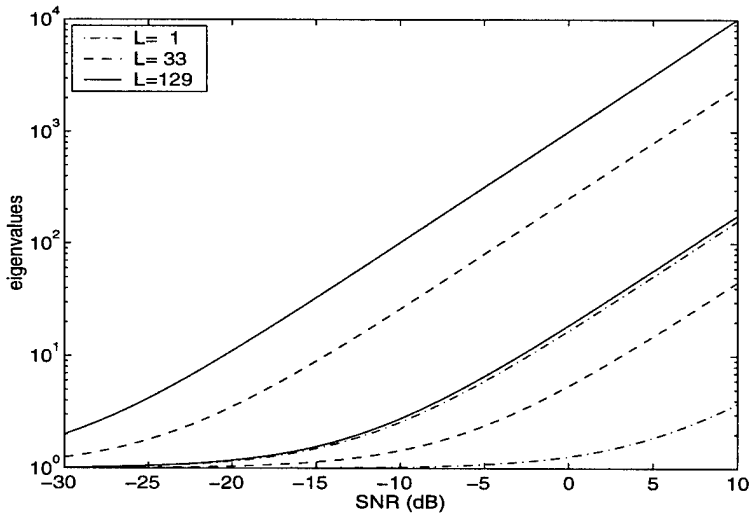
[22] G. Cunningham and W. Williams, "Fast implementation of generalized discrete time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 42, no. 6, pp. 1496–1508, June 1994.

[23] M. Kaveh and A. J. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-34, no. 2, pp. 331–340, April 1986.

[24] J. H. Wilkinson, *The Algebraic Eigenvalue Problem.* New York: Oxford University Press, 1965.

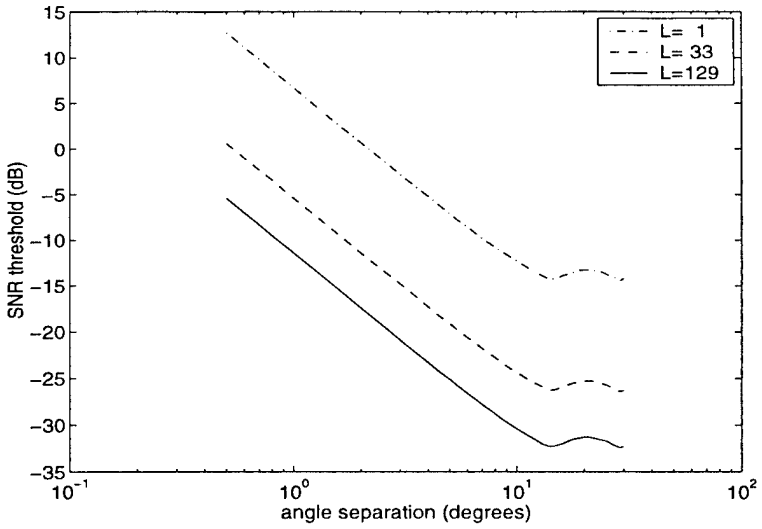


(a) $\theta_1 = -10^o, \theta_2 = 10^o$

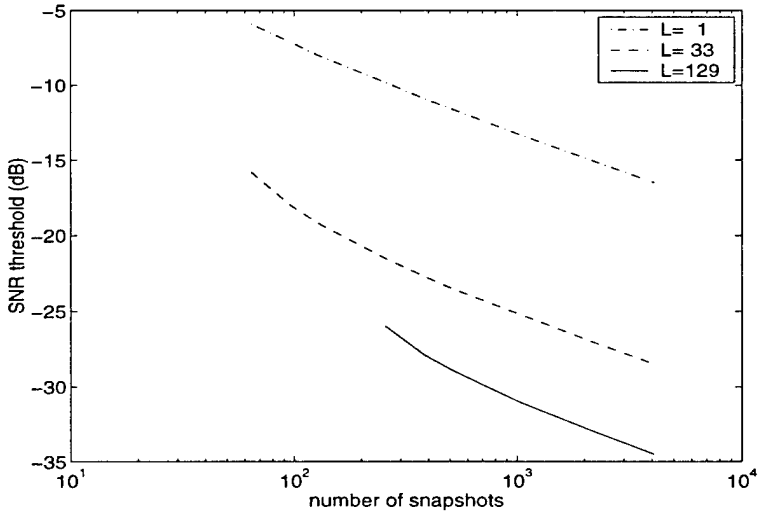


(b) $\theta_1 = -1^o, \theta_2 = 1^o$

Fig.1 The principal eigenvalues of correlation matrix and STFD matrix.



(a) SNR threshold vs. angle separation $N = 1024$



(b) SNR threshold vs. number of snapshots $\Delta\theta = 20°$

Fig.2 SNR thresholds to identify two signals ($m = 8$).

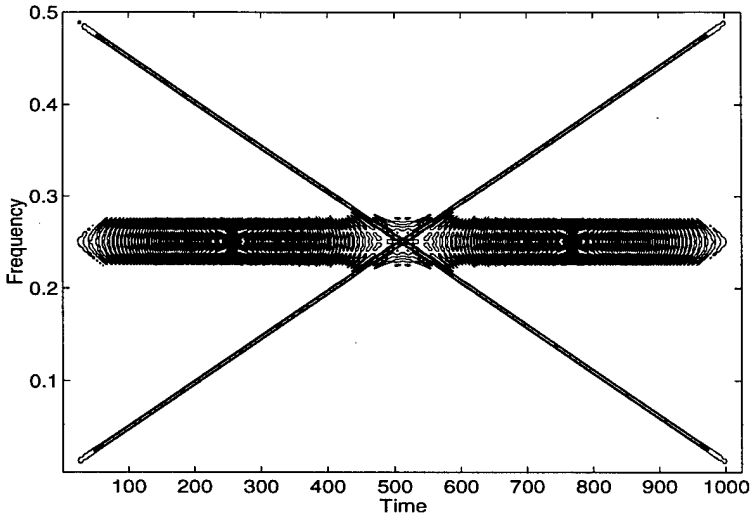


Fig.3 Pseudo Wigner-Ville distribution of the mixture of the two signals.

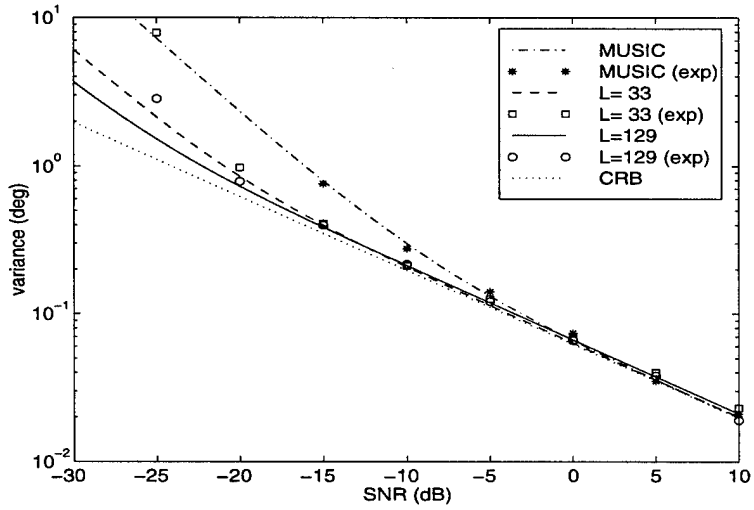


Fig.4 Variance of DOA estimation vs. SNR.

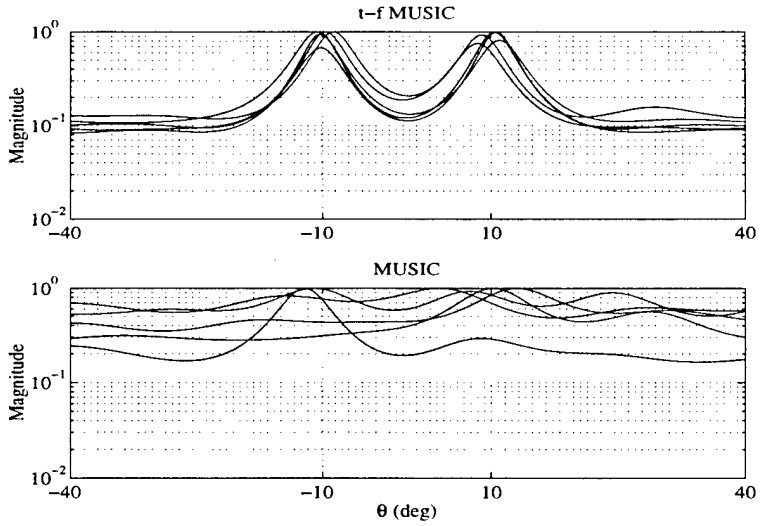


Fig.5 Estimated spatial spectra
(m=8, N=1024, SNR= -20 dB, $L = 129$ for t-f MUSIC).

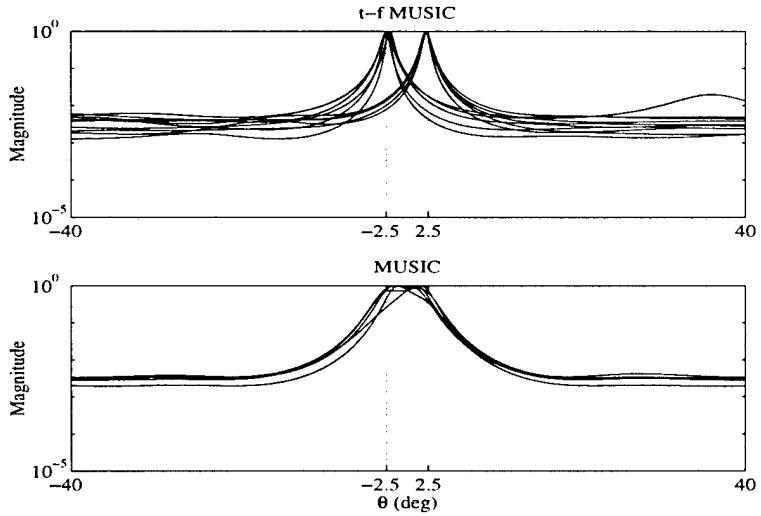


Fig.6 Estimated spatial spectra for closely spaced signals
(m=8, N=1024, SNR= -5 dB, $L = 129$ for t-f MUSIC).

Spatial Averaging of Time-Frequency Distributions for Signal Recovery in Uniform Linear Arrays

Yimin Zhang and Moeness G. Amin
Department of Electrical and Computer Engineering,
Villanova University, Villanova, PA 19085, USA

Abstract

This paper presents a new approach based on spatial time-frequency averaging for separating signals received by a uniform linear antenna array. In this approach, spatial averaging of the time-frequency distributions (TFDs) of the sensor data is performed at multiple time-frequency points. This averaging restores the diagonal structure of the source TFD matrix necessary for source separation. With spatial averaging, crossterms move from their off-diagonal positions in the source TFD matrix to become part of the matrix diagonal entries. It is shown that the proposed approach yields improved performance over the case when no spatial averaging is performed. Further, we demonstrate that, in the context of source separation, the spatially-averaged Wigner-Ville distribution outperforms the combined spatial-time-frequency averaged distributions, such as the one obtained by using the Choi-Williams distribution. Simulation examples involving the separation of two sources with close AM and FM modulations are presented.

This work is supported by Office of Naval Research under Grant No. N00014-98-1-0176.

I. Introduction

Recently, time-frequency distributions (TFD) have been employed for direction finding and blind source separation problems in sensor array processing [1-5]. The spatial time-frequency distributions (STFDs) were introduced in [1] and represented by a spatial matrix whose elements are the auto- and cross-time-frequency distributions of the data received at the different array sensors. STFD techniques are most appropriate to handle sources of nonstationary waveforms that are localized in the time-frequency domain. The robustness of the subspace estimates using STFD matrices is analyzed in [17], and shown to have an advantage over those obtained from the covariance matrices.

The application of STFDs to separating sources with distinct time-frequency (t-f) signatures is presented in [2]. In this reference, it is shown that the source TFD matrix, whose elements are the auto- and cross-TFD of the source signals, and the sensor data STFD have the same relationship as the one between the source and the data correlation matrices. This relationship is defined by the mixing, or the array manifold matrix. The steps applied in blind source separation based on second order statistics (the SOBI technique) outlined in [9] could therefore be used in the time-frequency formulation of the problem. The general theory of solving blind source separation problems using spatial arbitrary joint-variable distributions, including those of time and frequency, is given in [3]. In [4], the two arbitrary variables are chosen as the time-lag and frequency-lag, and the source separation was performed using spatial ambiguity functions. The use of STFDs for direction finding is discussed in [5] and [18], where the time-frequency MUSIC and the time-frequency maximum likelihood techniques are proposed.

Although blind source separations based on time-frequency distribution outperform the SOBI method for nonstationary signals, the fundamental problem with the bilinear time-frequency approach remains the need for the incorporation of STFD matrices computed only at the source autoterm points. Crossterms impede performance, as they reside on the off-diagonal elements of the source TFD matrix, and as such, violate its diagonal structure necessary for source separation. Identification of autoterm regions are often difficult for a large class of multi-component nonstationary signals, and even if properly identified, due to the complexity of the impinging signal time-frequency signatures and the use of finite

data records, autoterm regions cannot be entirely free from crossterm mainlobe or/and sidelobe contamination.

In this paper, we discuss the role of TFD crossterms and demonstrate the effect of spatial averaging on STFDs. By utilizing the Vandermonde structure of the array manifold matrix and performing spatial averaging on the spatial time-frequency distribution matrices, we set the off-diagonal elements of the corresponding source TFD matrix to zero. This is achieved by moving the crossterms from their off-diagonal positions to join the autoterms as diagonal entries of the source TFD matrix at one time-frequency point. In this respect, the performance of the source separation technique becomes much less dependent on the selection of the time-frequency points at which the STFD matrices are computed. It is shown that the spatially-averaged STFDs outperforms the case where no spatial averaging is performed, even when only autoterm points are involved in both cases.

Spatial averaging is a simple and well-known technique in conventional array processing [6]. It employs additional array sensors to reduce cross-correlation in coherent and correlated signal environments, and thereby permits proper angle-of-arrival (AOA) estimations and source separations. In this paper, we show that spatial averaging plays a key role in the underlying TFD-based source separation problem and its application leads to matrix diagonalization and crossterm mitigation. Spatial averaging gives robustness to time-frequency point selections and yields improved performance over other TFD-based techniques, specifically for sources whose signatures are closely separated in the time-frequency domain.

The restoration of the diagonal structure of the source TFD is only part of the problem. Source separation using spatially-averaged TFD evaluated at a single time-frequency point can still lead to noisy and non-unique results. Since the power distribution of the signals impinging on the array varies over the time-frequency plane, then different time-frequency points may exhibit different SNRs. The main two advantages of incorporating several spatially-averaged TFD matrices evaluated at different time-frequency points into a joint-diagonalization scheme are to avoid having degenerate eigenvalues and to reduce the possibility of choosing a point with high noise contamination.

It is noted that, unlike the method in [2], the proposed approach requires the information

on the array manifold, and is sensitive to the calibration error. In this case, conventional AOA estimation methods, such as the maximum likelihood [11], matrix pencil [12], MUSIC [13], root MUSIC [14, 15], and ESPRIT [16] techniques, can also be used to estimate the mixing matrix, and further to separate the source signals. The proposed approach not only requires no angular search but also enjoys the discriminatory property of TFD-based array processing, where fewer sources can be considered by only selecting their respective time-frequency signatures [17–19].

This paper is organized as follows. In Section II, the source separation approach based on spatial time-frequency distribution is briefly summarized. In Section III, we introduce the spatially averaged time-frequency distributions, and discuss the difference between spatial averaging and kernel methods in handling the crossterm problem. Simulation results demonstrating the usefulness of the proposed technique are given in Section IV.

II. Source Separation Based on Spatial Time-Frequency Distributions

A. Spatial Time-Frequency Distributions

In many practical situations, the data vector $\mathbf{x}(t)$ for an N-element array follows an instantaneous mixture model and is given by

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \tag{1}$$

where $\mathbf{x}(t) = [x_0(t), \cdots, x_{N-1}(t)]^T$ is the data snapshot vector at time t, and the superscript ^{T} denotes transpose. The vector $\mathbf{s}(t) = [s_1(t), \cdots, s_n(t)]^T$ contains n source signals at the same time, and $\mathbf{n}(t)$ is the additive noise vector. This model is commonly used in the field of narrowband array processing. The vector $\mathbf{y}(t) = [y_0(t), \cdots, y_{N-1}(t)]^T$ contains the noise-free array output. The mixing matrix $\mathbf{A}$ is the transfer function between the source signals and the data at the array sensors. We assume that the mixing matrix $\mathbf{A}$ is full column rank.

The source signal vector $\mathbf{s}(t)$ is assumed to be a deterministic signal vector with correlation matrix

$$\mathbf{R}_{\mathbf{ss}}(\tau) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbf{s}(t+\tau)\mathbf{s}^H(t) \tag{2}$$

where superscript ^{H} denotes the conjugate transpose of a matrix or a vector. In reference [2], it is assumed that $\mathbf{R}_{\mathbf{ss}}(\tau) = diag[r_{11}(\tau), \cdots, r_{nn}(\tau)]$, where $diag[\cdot]$ is the diagonal

matrix formed with the elements of its vector valued argument, and $r_{ii}(\tau) = \lim_{T\to\infty} \sum_{t=1}^{T} s_i(t+\tau)s_i^*(t)$ denotes the correlation of $s_i(t)$. This assumption implies that the components $s_i(t)$, $1 \leq i \leq n$, are mutually uncorrelated. However, in our proposed method, this assumption is no longer necessary.

The additive noise $\mathbf{n}(t)$ is modeled as a stationary, temporally white, zero-mean complex random process independent of the source signals. For simplicity, we also require $\mathbf{n}(t)$ to be spatially white, i.e.,

$$E\left[\mathbf{n}(t+\tau)\mathbf{n}^H(t)\right] = \sigma_n^2\delta(\tau)\mathbf{I} \tag{3}$$

where $\delta(\tau)$ is the Kronecker delta and $\mathbf{I}$ denotes the identity matrix. Since the signal power and the signal ordering are indeterminable in source separations [3], we simplify the problem by treating the source signals as if they have unit power. Accordingly

$$\mathbf{R_{ss}}(0) = \mathbf{I} \text{ and } \mathbf{R_{yy}} = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}\mathbf{y}(t)\mathbf{y}^H(t) = \mathbf{A}\mathbf{A}^H. \tag{4}$$

The discrete-time form of Cohen's class of TFD for signal $x(t)$ is given by [7]

$$D_{xx}(t,f) = \sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\phi(m,l)x(t+m+l)x^*(t+m-l)e^{-4\pi fl} \tag{5}$$

where t and f represent the time index and the frequency index, respectively. The kernel $\phi(m,l)$ characterizes the TFD and is a function of both the time and lag variables. The cross-TFD of two signals $x_i(t)$ and $x_j(t)$ is defined by [7]

$$D_{x_ix_j}(t,f) = \sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\phi(m,l)x_i(t+m+l)x_j^*(t+m-l)e^{-4\pi fl}. \tag{6}$$

One possible definition of spatial time-frequency distribution (STFD) is given in [2] and incorporates both equations (5) and (6),

$$\mathbf{D_{xx}}(t,f) = \sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\phi(m,l)\mathbf{x}(t+m+l)\mathbf{x}^H(t+m-l)e^{-4\pi fl} \tag{7}$$

where $[\mathbf{D_{xx}}(t,f)]_{i,j} = D_{x_ix_j}(t,f)$, for $i,j = 0,\cdots,N-1$. It is shown in the next section that other forms of STFD can be more useful in the context of source separation. Under the linear data model of Eq. (1), and assuming noise-free environment, the STFD matrix in (7) takes the following simple structure

$$\mathbf{D_{xx}} = \mathbf{A}\mathbf{D_{ss}}(t,f)\mathbf{A}^H \tag{8}$$

where $\mathbf{D}_{\mathbf{ss}}(t,f)$ is the signal TFD matrix whose entries are the auto- and cross-TFDs of the sources. Eq. (8) is similar to the formula that is commonly used in conventional blind source separation and direction-of-arrival (DOA) estimation problems [8,9], relating the signal correlation matrix to the data spatial correlation matrix. If $\mathbf{D}_{\mathbf{ss}}(t,f)$ is a full-rank matrix, the two subspaces spanned by the principle eigenvectors of $\mathbf{D}_{\mathbf{xx}}(t,f)$ and the columns of $\mathbf{A}$ become identical. In this case, direction finding techniques based on eigenstructures can be applied. If $\mathbf{D}_{\mathbf{ss}}(t,f)$ is diagonal, i.e., the signal cross-TFDs at the time-frequency point (t,f) are zeros, the mixture matrix and the signal waveforms can be recovered using blind source separation methods [1,2].

B. Source Separation Based on Spatial Time-Frequency Distributions

The source separation algorithm based on spatial time-frequency distributions is an essential part of the proposed method. The algorithm is given in details in reference [2] and is summarized below.

The first step is the whitening of the signal part $\mathbf{y}(t)$ of the observation. This is achieved by applying a whitening matrix $\mathbf{W}$ to $\mathbf{y}(t)$, i.e., an $n \times N$ matrix satisfying:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbf{W}\mathbf{y}(t)\mathbf{y}^H(t)\mathbf{W}^H = \mathbf{W}\mathbf{R}_{\mathbf{yy}}\mathbf{W}^H = \mathbf{W}\mathbf{A}\mathbf{A}^H\mathbf{A}^H = \mathbf{I}. \tag{9}$$

$\mathbf{WA}$ is an $n \times n$ unitary matrix $\mathbf{U}$, and matrix $\mathbf{A}$ can be written as

$$\mathbf{A} = \mathbf{W}^{\#}\mathbf{U} \tag{10}$$

where superscript $^{\#}$ denotes pseudo-inverse. The whitened process $\mathbf{z}(t) = \mathbf{W}\mathbf{x}(t)$ still obeys a linear model,

$$\mathbf{z}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}\left[\mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)\right] = \mathbf{U}\mathbf{s}(t) + \mathbf{W}\mathbf{n}(t). \tag{11}$$

By pre- and post-multiplying the STFD matrices $\mathbf{D}_{\mathbf{xx}}(t,f)$ by $\mathbf{W}$, we obtain

$$\mathbf{D}_{\mathbf{zz}}(t,f) = \mathbf{W}\mathbf{D}_{\mathbf{xx}}(t,f)\mathbf{W}^H \tag{12}$$

which is, in essence, the STFD of the whitened data vector $\mathbf{z}(t)$. From the definitions of $\mathbf{W}$ and $\mathbf{U}$,

$$\mathbf{D}_{\mathbf{zz}}(t,f) = \mathbf{U}\mathbf{D}_{\mathbf{ss}}(t,f)\mathbf{U}^H. \tag{13}$$

Equation (13) shows that if $\mathbf{D_{ss}}(t,f)$ is diagonal, then any whitened data STFD matrix is diagonal in the basis of the columns of the matrix $\mathbf{U}$, and the eigenvalues of $\mathbf{D_{zz}}(t,f)$ are the diagonal entries of $\mathbf{D_{ss}}(t,f)$. An estimate $\hat{\mathbf{U}}$ of the unitary matrix $\mathbf{U}$ may be obtained as a signal subspace of a whitened STFD matrix evaluated at a time-frequency point corresponding to the signal autoterm. The source signals can then be estimated as $\hat{\mathbf{s}}(t) = \hat{\mathbf{U}}\hat{\mathbf{W}}\mathbf{x}(t)$, and the mixing matrix $\mathbf{A}$ is estimated by $\hat{\mathbf{A}} = \hat{\mathbf{W}}^{\#}\hat{\mathbf{U}}$.

Although the unitary matrix can be obtained from a single time-frequency point, STFDs corresponding to different (t,f) points should be incorporated, so as to reduce the possibility of having degenerate eigenvalues and subsequently non-unique solutions. The joint-diagonalization (JD) scheme can be used to incorporate multiple time-frequency points [2]. This scheme forms K STFD matrices $\{\mathbf{D_{zz}}(t_i,f_i)|i=1,\cdots,K\}$ at a set of preferable K time-frequency autoterm points. The unitary matrix $\hat{\mathbf{U}}$ is then obtained as the joint diagonalizer of the set $\{\mathbf{D_{zz}}(t_i,f_i)|i=1,\cdots,K\}$.

III. Spatial Averaging Time-Frequency Distributions

A. Spatial Averaging Methods

Spatial averaging method was introduced by Pillai [6] to restore the full-rank property of the signal correlation matrix in the presence of coherent signals. Unlike other spatial smoothing methods [20–23], which only restore the full rank property of the mixing matrix when the impinging signals are coherent, the spatial averaging method enforces the diagonal structure of the signal correlation matrix. This diagonal matrix property is essential to perform source separation, as previously discussed. Here, we present the role of spatial averaging in the context of TFD analysis, and propose signal separation using joint diagonalization based on spatial averaging of spatial TFD matrices.

The basic idea of spatial averaging is to use subarrays of a uniform linear array to obtain an averaged correlation matrix, or in the underlying problem, an averaged STFD matrix, with the off-diagonal elements set to zero.

Without loss of generality, we consider a simple example of $n = 2$, i.e., there are only two sources, $s_1(t)$ and $s_2(t)$. The result is generally true for n sources and N sensors, as long as $n < N$.

By ignoring the effect of noise, the received signal at the i-th array sensor ($i = 0, 1, \cdots, N-1$) is represented by

$$x_i(t) = x_i^{(1)}(t) + x_i^{(2)}(t) = s_1(t)e^{-jd_i\omega_1} + s_2(t)e^{-jd_i\omega_2} \tag{14}$$

where $\omega_k = 2\pi \sin\phi_k/\lambda, k = 1, 2$, are the spatial radian frequencies, ϕ_k are the angles-of-arrival, λ is the RF wavelength, and d_i is the distance between the 0-th and the i-th array sensors. The cross-TFD of $x_i(t)$ and $x_j(t)$, assuming uniform linear array, is

$$\begin{aligned} D_{x_ix_j}(t,f) &= D_{x_i^{(1)}x_j^{(1)}}(t,f) + D_{x_i^{(2)}x_j^{(1)}}(t,f) + D_{x_i^{(2)}x_j^{(2)}}(t,f) + D_{x_i^{(1)}x_j^{(2)}}(t,f) \\ &= \left[D_{s_1s_1}(t,f) + D_{s_2s_1}(t,f)e^{-jd_i(\omega_2-\omega_1)}\right] e^{-j(d_i-d_j)\omega_1} \\ &\quad + \left[D_{s_2s_2}(t,f) + D_{s_1s_2}(t,f)e^{-jd_i(\omega_2-\omega_1)}\right] e^{-j(d_i-d_j)\omega_2}. \end{aligned} \tag{15}$$

Due to the presence of the cross-terms (second term in each bracket in (15)), the TFD matrix $\mathbf{D_{xx}}(t,f)$ does not provide the proper information to carry out source separations.

The auto- and cross-TFD of the data $x_0(t)$ and $x_i(t)$, $i = 0, 1, \cdots, N-1$, is

$$D_{x_0x_i}(t,f) = \left[D_{s_1s_1}(t,f) + D_{s_2s_1}(t,f)\right] e^{jd_i\omega_1} + \left[D_{s_2s_2}(t,f) + D_{s_1s_2}(t,f)\right] e^{jd_i\omega_2} \tag{16}$$

where we used the sensor receiving $x_0(t)$ as the reference sensor and set $d_0 = 0$. Denote $b_1(t,f) = D_{s_1s_1}(t,f) + D_{s_2s_1}(t,f)$ and $b_2(t,f) = D_{s_2s_2}(t,f) + D_{s_1s_2}(t,f)$. The values of $b_1(t,f)$ and $b_2(t,f)$ are generally complex. If $b_1(t,f)$ and $b_2(t,f)$ are real, then the Hermitian Toeplitz spatial time-frequency matrix

$$\overline{\mathbf{D}}_{\mathbf{xx}}(t,f) = \begin{bmatrix} D_{x_0x_0}(t,f) & D_{x_0x_1}(t,f) & \cdots & D_{x_0x_{N-1}}(t,f) \\ D^*_{x_0x_1}(t,f) & D_{x_0x_0}(t,f) & \cdots & D_{x_0x_{N-2}}(t,f) \\ \vdots & \vdots & \ddots & \vdots \\ D^*_{x_0x_{N-1}}(t,f) & D^*_{x_0x_{N-2}}(t,f) & \cdots & D_{x_0x_0}(t,f) \end{bmatrix} \tag{17}$$

generated from the cross-TFDs $D_{x_0x_0}(t,f), D_{x_0x_1}(t,f), \cdots, D_{x_0x_{N-1}}(t,f)$ between the data samples at the reference sensor and those at other sensors of the array can be expressed as [24]

$$\overline{\mathbf{D}}_{\mathbf{xx}}(t,f) = \mathbf{A}\overline{\mathbf{D}}_{\mathbf{ss}}(t,f)\mathbf{A}^H \tag{18}$$

where $\mathbf{A}$ is a Vandermonde matrix, and

$$\overline{\mathbf{D}}_{\mathbf{ss}}(t,f) = diag\left[b_1(t,f),\ b_2(t,f)\right] \tag{19}$$

is the corresponding source TFD matrix. Note that $\overline{\mathbf{D}}_{\mathbf{xx}}(t,f)$ has a different structure from that of the STFD matrix defined in (7), and was used in reference [2] for blind source separation. Clearly, (18) has the same form as (8), but $\overline{\mathbf{D}}_{\mathbf{ss}}(t,f)$ here is diagonal, even if the selected (t,f) point corresponds to a crossterm.

In the case of complex signal waveforms, the realness and the diagonal structure of $\overline{\mathbf{D}}_{\mathbf{ss}}(t,f)$ can be restored by spatial averaging. We add $N-1$ array sensors symmetrically about the reference point, as shown in Fig. 1. The received signal at i-th sensor of the new set is,

$$x_{-i}(t) = x_{-i}^{(1)}(t) + x_{-i}^{(2)}(t) = s_1(t)e^{jd_i\omega_1} + s_2(t)e^{jd_i\omega_2}. \tag{20}$$

The new cross-TFD of $x_0(t)$ and $x_{-i}(t)$ is,

$$D_{x_0x_{-i}}(t,f) = \left[D_{s_1s_1}(t,f) + D_{s_2s_1}(t,f)\right]e^{-jd_i\omega_1} + \left[D_{s_2s_2}(t,f) + D_{s_1s_2}(t,f)\right]e^{-jd_i\omega_2}. \tag{21}$$

The spatial averaging of (16) and (21) is given by

$$\tilde{D}_{xx}^{(i)}(t,f) = \left[D_{x_0x_i}(t,f) + D_{x_0x_{-i}}^{*}(t,f)\right]/2 = c_1(t,f)e^{jd_i\omega_1} + c_2(t,f)e^{jd_i\omega_2} \tag{22}$$

where

$$c_1(t,f) = D_{s_1s_1}(t,f) + Re\left[D_{s_2s_1}(t,f)\right], \quad c_2(t,f) = D_{s_2s_2}(t,f) + Re\left[D_{s_1s_2}(t,f)\right].$$

Since the terms in the brackets in (21) are all real, the matrix formed from the TFDs (22)

$$\tilde{\mathbf{D}}_{\mathbf{xx}}(t,f) = \begin{bmatrix} \tilde{D}_{xx}^{(0)}(t,f) & \tilde{D}_{xx}^{(1)}(t,f) & \cdots & \tilde{D}_{xx}^{(N-1)}(t,f) \\ \tilde{D}_{xx}^{(1)*}(t,f) & \tilde{D}_{xx}^{(0)}(t,f) & \cdots & \tilde{D}_{xx}^{(N-2)}(t,f) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{D}_{xx}^{(N-1)*}(t,f) & \tilde{D}_{xx}^{(N-2)*}(t,f) & \cdots & \tilde{D}_{xx}^{(0)}(t,f) \end{bmatrix} \tag{23}$$

is Hermitian and Toeplitz. This matrix is referred to as the spatially-averaged TFD (SATFD) matrix. Similar to the real TFD case, in the noise-free environment, the SATFD matrix in (23) can be expressed as

$$\tilde{\mathbf{D}}_{\mathbf{xx}}(t,f) = \mathbf{A}\tilde{\mathbf{D}}_{\mathbf{ss}}(t,f)\mathbf{A}^H \tag{24}$$

where

$$\tilde{\mathbf{D}}_{\mathbf{ss}}(t,f) = diag\left[c_1(t,f),\ c_2(t,f)\right]. \tag{25}$$

The off-diagonal elements of $\tilde{\mathbf{D}}_{\mathbf{ss}}(t,f)$ are zero, where as the matrix diagonal entries are now made up of both autoterms and crossterms of the impinging source signals. By enforcing the diagonal structure of the source TFD matrix $\tilde{\mathbf{D}}_{\mathbf{ss}}(t,f)$, spatial averaging of the Hermitian Toeplitz STFD matrices extends the validity of the TFD-based signal separation in the presence of cross-TFDs.

The steps for source separation used in [2] and summarized in Section II can be applied to the SATFD $\tilde{\mathbf{D}}_{\mathbf{xx}}(t,f)$ instead of the STFD $\mathbf{D}_{\mathbf{xx}}(t,f)$. With spatial averaging, the incorporation of STFDs at only autoterm points into the joint-diagonalization scheme is no longer crucial to achieve good performance.

B. Comparison between Spatial Averaging and Kernel Methods

There are two sources of crossterms in the underlying source separation problem. The first type are the crossterms that are the results of the interactions between the components of the same source signal. Whether we use the STFD defined in (7) or in (17), those crossterms are not harmful to the blind source separation problem, since they always reside, along with the autoterms, on the main diagonal of the source TFD matrix. The other type of crossterms are those generated from the interactions between two signal components belonging to two different sources. These crossterms are associated with cross-TFDs of the source signals and, at any given time-frequency point, they constitute the off-diagonal entries of the source TFD matrices. The crossterms generated from the data cross-TFDs violate the basic assumption in the problem of source separation regarding the diagonal structure of the source TFD matrix. We must therefore select the t-f points that belong to autoterm regions where crossterm contributions are at minimum. However, the selection of autoterm points is often difficult in the absence of *a priori* information of the source signals, specifically for low SNR or when the signals have highly overlapping time-frequency signatures. The later case can be encountered in radar echoes and acoustic signal processing.

The use of smoothing time-frequency kernel for crossterm reduction is a candidate solution of the above problem. The main function of this kernel in the context of source separation is to prevent the selection and incorporation of crossterm points in the joint-diagonalization scheme, as well as to reduce the contribution of crossterms at selected

autoterm points. In essence, the fundamental role of the t-f kernel is to make the source TFD matrices as close to a diagonal structure as possible. The t-f kernel can be applied to both forms of STFDs in (7) and (23). It is noteworthy that the smoothing kernel does not distinguish between the aforementioned two types of crossterms, and accordingly it reduces all entries of the source TFD matrix, including the diagonal elements.

The problem with the smoothing kernel is fourfold. First, for sources with closely separated time-frequency signatures, the effectiveness of the smoothing kernel in reducing crossterms is highly impaired. Second, reduction of crossterms depends on their time-frequency locations, especially when fixed shape kernels are used. For example, t-f kernels satisfying the marginal properties are not suitable for removing the crossterms which lie on the time-lag and frequency-lag axes in the ambiguity domain. Third, depending on the employed t-f kernel, part or all of the crossterms may be displaced to mount on the selected autoterm points. The situation can make the source TFD matrix to further deviate from a diagonal structure, cause performance deterioration from the case when no smoothing is applied. We refer to this undesired property as the smoothing problem. Fourth, since source separation is often performed incorporating a finite number of data samples, the intrusion of crossterms on autoterm regions cannot be prevented or entirely removed. This is because the window spreads out the crossterms in the time-frequency domain so that the mainlobe or/and the sidelobes of the crossterms are deemed to overlap with the signal autoterms. We refer to this undesired property as the leakage problem in STFDs. In addition to the above drawbacks, the t-f kernel ignores the fact that the first type of crossterms need not be smoothed, as its appearance along the diagonal elements can improve the effective signal to noise ratio.

The spatial averaging of the STFD defined in (23) at a given (t, f) point does not smooth or reduce the crossterms at that point, but rather move them from their residence on the off-diagonal matrix entries to be part of the matrix diagonal elements. The other part represents the contribution of the autoterms at the same point. Therefore, not only we are able to set the off-diagonal elements of the source TFD matrix to zeros, but also we can improve performance by selecting the (t, f) points of peak values, irrespective of whether these points belong to autoterm or crossterm regions.

IV. Performance Evaluation

A. Performance Index

We use a slightly modified version of the performance index applied in [2] to evaluate the performance of the proposed source separation technique. The estimate of the source signal vector is computed by applying the pseudo-inverse of the estimated mixing matrix $\hat{\mathbf{A}}$ to the received signal vector $\mathbf{x}(t)$, i.e.,

$$\hat{\mathbf{s}}(t) = \hat{\mathbf{A}}^{\#}\mathbf{x}(t) = \hat{\mathbf{A}}^{\#}\mathbf{A}\mathbf{s}(t) + \hat{\mathbf{A}}^{\#}\mathbf{n}(t) \tag{26}$$

where $\hat{\mathbf{A}} = \hat{\mathbf{W}}^{\#}\hat{\mathbf{U}}$. We stress that in general, this procedure is not optimal for recovering the source signals based on an estimate $\hat{\mathbf{A}}$. For large enough sample size, matrix $\hat{\mathbf{A}}$ should be close to the true one $\mathbf{A}$, so that $\hat{\mathbf{A}}^{\#}\mathbf{A}$ well approximates the identity matrix. We normalize matrix $\hat{\mathbf{A}}$ by

$$\hat{\mathbf{A}}_e = \hat{\mathbf{A}} diagonal\left(\hat{\mathbf{A}}^{\#}\mathbf{A}\right) \tag{27}$$

where $diagonal(\mathbf{F})$ denotes the matrix formed by the diagonal elements of $\mathbf{F}$. As such, the diagonal elements of $\hat{\mathbf{A}}_e^{\#}\mathbf{A}$ become exactly one, giving more meaning to the performance index

$$I_{pq} = E\left|\left(\hat{\mathbf{A}}_e^{\#}\mathbf{A}\right)_{pq}\right|^2 \tag{28}$$

which defines the interference-to-signal ratio (ISR). Thus, I_{pq} measures the ratio of the power of the interference of q-th source signal to the power of the p-th source signal. As the measure of the global quality of the separation process, we also apply the global rejection level to evaluate the overall performance of the proposed method

$$I_{perf} = \sum_{q \neq p} I_{pq}. \tag{29}$$

B. Effect of Crossterms between Source Signals

In this section, we examine the effect of the time-frequency crossterms on source separation performance when spatial averaging is not applied. To simplify the problem, we assume that $\mathbf{R}_{\mathbf{ss}}$ is an identity matrix. When crossterms are present at the off-diagonal elements of the TFD matrix $\mathbf{D}_{\mathbf{ss}}(t,f)$, then

$$\mathbf{D}_{\mathbf{ss}}(t,f) = \mathbf{P}(t,f)\mathbf{G}(t,f)\mathbf{P}^{H}(t,f) \tag{30}$$

where $\mathbf{G}(t,f)$ is the diagonal matrix with the eigenvalues at the diagonal elements, and $\mathbf{P}(t,f)$ is the matrix whose columns are the corresponding eigenvectors. Note that all the above matrices depend on the selected (t,f) point. Substituting (30) in (8), the STFD matrix of the data vector under noise-free conditions becomes

$$\mathbf{D_{xx}}(t,f) = \mathbf{A}\mathbf{D_{ss}}(t,f)\mathbf{A}^H = \mathbf{A}\mathbf{P}(t,f)\mathbf{G}(t,f)\mathbf{P}^H(t,f)\mathbf{A}^H \tag{31}$$

and the STFD matrix of the whitened array signal vector is

$$\mathbf{D_{zz}}(t,f) = \mathbf{W}\mathbf{A}\mathbf{D_{ss}}(t,f)\mathbf{A}^H\mathbf{W}^H = \mathbf{W}\mathbf{A}\mathbf{P}(t,f)\mathbf{G}(t,f)\mathbf{P}^H(t,f)\mathbf{A}^H\mathbf{W}^H. \tag{32}$$

Since $\mathbf{G}(t,f)$ is diagonal, $\mathbf{WAP}(t,f)$ is unitary. If the estimated mixing matrix $\hat{\mathbf{A}}$ is provided based on a single (t,f) point, then from (32),

$$\hat{\mathbf{A}} = \mathbf{W}^{\#}\mathbf{W}\mathbf{A}\mathbf{P}(t,f) = \mathbf{A}\mathbf{P}(t,f) \tag{33}$$

which is dependent on the unitary matrix $\mathbf{P}(t,f)$. Furthermore,

$$\hat{\mathbf{A}}^{\#}\mathbf{A} = [\mathbf{A}\mathbf{P}(t,f)]^{\#}\mathbf{A} = \mathbf{P}^H(t,f) \tag{34}$$

and

$$\begin{aligned}
\hat{\mathbf{A}}_e^{\#}\mathbf{A} &= \left[diagonal\left(\hat{\mathbf{A}}^{\#}\mathbf{A}\right)\right]^{-1}\hat{\mathbf{A}}^{\#}\mathbf{A} \\
&= \left[diagonal\left(\mathbf{P}^H(t,f)\right)\right]^{-1}\mathbf{P}^H(t,f) \\
&= \begin{bmatrix} p_{11}^{-1}(t,f) & & & \mathbf{O} \\ & p_{22}^{-1}(t,f) & & \\ & & \ddots & \\ \mathbf{O} & & & p_{nn}^{-1}(t,f) \end{bmatrix}^* \begin{bmatrix} p_{11}(t,f) & p_{21}(t,f) & \cdots & p_{n1}(t,f) \\ p_{12}(t,f) & p_{22}(t,f) & \cdots & p_{n2}(t,f) \\ \vdots & \vdots & \ddots & \vdots \\ p_{1n}(t,f) & p_{2n}(t,f) & \cdots & p_{nn}(t,f) \end{bmatrix}^* \\
&= \begin{bmatrix} 1 & p_{22}^{-1}p_{21}(t,f) & \cdots & p_{nn}^{-1}p_{n1}(t,f) \\ p_{11}^{-1}p_{12}(t,f) & 1 & \cdots & p_{nn}^{-1}p_{n2}(t,f) \\ \vdots & \vdots & \ddots & \vdots \\ p_{11}^{-1}p_{1n}(t,f) & p_{22}^{-1}p_{2n}(t,f) & \cdots & 1 \end{bmatrix}^*
\end{aligned} \tag{35}$$

where $p_{ij} = [\mathbf{P}(t,f)]_{ij}$. Accordingly, the performance index becomes

$$I_{pq} = \left|p_{qq}^{-1}(t,f)p_{qp}(t,f)\right|^2 \tag{36}$$

and the global rejection level is given by

$$I_{perf} = \sum_{q \neq q} I_{pq} = \sum_{q=1}^{n} |p_{qq}(t,f)|^{-2} \sum_{p=1 p \neq q}^{n} |p_{qp}(t,f)|^2 = \sum_{q=1}^{n} |p_{qq}(t,f)|^{-2} - n. \qquad (37)$$

In general, since the absolute values of $p_{qq}(t,f)$ are always equal to or smaller than 1, the global rejection level I_{perf} takes a positive value. It is clear that $I_{perf} = 0$ only when $p_{qq}(t,f) = 1$ holds true for all q. That is, $\mathbf{P}$ is an identity matrix, which implies that there is no off-diagonal non-zero elements in matrix $\mathbf{D_{ss}}(t,f)$.

Consider the specific case of $n = 2$. If we select a (t,f) point where the contributions of the two sources to the source TFD matrix are the same, i.e., $D_{s_1 s_1}(t,f) = D_{s_2 s_2}(t,f)$, and since $D_{s_1 s_2}(t,f) = D^*_{s_2 s_1}(t,f)$ by definition, then it is straightforward to show that $|p_{qq}(t,f)| = 1/\sqrt{2}$. In this case, I_{perf} is constant equal to 2. The (t,f) points having such property include all crossterms at which the autoterms have equal contributions.

C. Simulation Results

In this section we demonstrate the effectiveness of the spatially-averaged time-frequency distributions in source separations. The whitening joint-diagonalization scheme [2] is used for incorporating multiple time-frequency points into the proposed spatial averaging method. In all simulations, two sources with the chirp signals

$$s_1(t) = e^{-j\mu\frac{t^2}{2}}, \quad s_1(t) = e^{-j\mu\frac{t^2}{2}-(\alpha+j\omega)t} \qquad (38)$$

are used, where μ is chosen as 0.008π. Different values of ω and α are considered. These values control the frequency offset and amplitude variation between the two signals and can be chosen to yield closely or widely separated source signatures in the time-frequency domain. 128 data samples are considered, from which a time-frequency matrix of 128×128 is formed. The DOAs of the two signals $s_1(t)$ and $s_2(t)$ are set equal to 30° and 60°, respectively, from the broadside direction. Furthermore, we assume an equi-spaced 5-element linear array with the interelement spacing 0.5λ, where λ is the wavelength. Subsequently, when the spatial averaging method is used, two sub-arrays are formed, each with 3 elements.

In the first set of simulations, we choose α=0, i.e., neither signal is amplitude modulated. The Wigner-Ville (WV) distribution of each signal is shown in Fig. 2, where δf $(=\omega/2\pi)$ is

set equal to 0.05. Fig. 3 shows the time-frequency distribution of the mixed signals at the center array sensor. No noise is present for this case. It is clear that the crossterms lie in the middle of the two chirps, and their amplitude change periodically. Fig. 4(a) shows the time-frequency distributions of the separated signals using the technique in [2], where joint diagonalization is used without the utilization of the proposed spatial averaging method. Three time-frequency (t, f) points are used at $t = 32$, 64, and 96. The frequency f is chosen so that the TFD computed using the data at the center array sensor is the largest at each t. Peak values of the WV distribution may either correspond to autoterms or crossterms. In this case, out of three (t, f) points, one crossterm point and two autoterm peaks were selected. The obtained $\hat{\mathbf{A}}_e^{\#}\mathbf{A}$ matrix is

$$\hat{\mathbf{A}}_e^{\#}\mathbf{A} = \begin{bmatrix} 1.00 + j0.00 & 0.19 + j0.65 \\ -0.21 + j0.63 & 1.00 + j0.00 \end{bmatrix}$$

and the computed global rejection level I_{perf} is -0.43 dB. The result is clearly unsatisfactory, as the matrix $\hat{\mathbf{A}}_e^{\#}\mathbf{A}$ is far from the identity matrix and crossterms appear in the separated signals.

Next, we force the selection of autoterm peaks by only considering the (t, f) points along the instantaneous frequencies of the two input signals at the same above time instants. Although no crossterm point is selected, yet as discussed in Section III, because of the finite data record, the crossterms leak into autoterm regions, causing the source TFD matrix to deviate from a diagonal structure. We show in Fig. 4(b) the result of source separation when only the autoterm points are considered. The obtained $\hat{\mathbf{A}}_e^{\#}\mathbf{A}$ matrix becomes

$$\hat{\mathbf{A}}_e^{\#}\mathbf{A} = \begin{bmatrix} 1.00 + j0.00 & 0.00 - j0.06 \\ -0.01 - j0.01 & 1.00 + j0.00 \end{bmatrix}$$

and the respective global rejection level I_{perf} is -23.96 dB. It is clear that the source separation performance is greatly improved. This good performance implies that the contributions of crossterms at the three selected autoterm points were insignificant, implying that the corresponding source TFD matrices in this case were close to diagonal.

Fig. 5 shows the time-frequency distributions of the separated signals at the same condition as Fig. 4(a), except with the utilization of the proposed spatial averaging method. Spatial averaging entirely mitigates the effect of crossterms. It is clear that the time-frequency distributions of the separated signals are the same as those of the original source

signals, and $\hat{\mathbf{A}}_e^{\#}\mathbf{A}$ are exactly identity matrices. Similar results can be obtained when all three (t, f) points are autoterms.

Fig. 6 shows the global rejection level I_{perf} versus the frequency difference δf between the two chirps, where the input SNR is 20 dB. When the proposed spatial averaging method is used, the global rejection level maintains low values. On the other hand, without spatial averaging, the global rejection levels become very high. The main reason of the large fluctuation of the I_{perf} without spatial averaging is that the influence as well as the number of crossterm points incorporated in the joint-diagonalization scheme varies with the frequency difference δf (when δf=0.1, no crossterm points were selected). Note that the crossterms of the Wigner-Ville distribution remain high even when the frequency difference is large. When selected, these terms put large values along the off-diagonal terms of the source TFD matrix, and subsequently cause considerable error, as evident from the figure. However, when only autoterm (t, f) points are used, the global rejection level decreases as δf increases. In this case, the matrix off-diagonal elements are the crossterm values at the autoterm points which become smaller for higher values of δf.

Next we show the effect of using time-frequency smoothing kernels for reduced interference terms. The Choi-Williams (CW) distribution [10] is considered with σ=1. Fig. 7 shows the CW distribution of each signal separately, whereas the CW distribution of the mixed signals at the center array sensor is depicted in Fig. 8. The signals are the same as the ones used in the WV distribution simulations with δf=0.05. Fig.9(a) shows the CW distributions of the separated signals. The obtained $\hat{\mathbf{A}}_e^{\#}\mathbf{A}$ matrix is

$$\hat{\mathbf{A}}_e^{\#}\mathbf{A} = \begin{bmatrix} 1.00 + j0.00 & 0.03 + j0.70 \\ -0.05 + j0.67 & 1.00 + j0.00 \end{bmatrix}$$

and the respective global rejection level I_{perf} is -0.26 dB. At this small frequency offset, effective smoothing of crossterms is difficult, and as a result, even with the use of time-frequency kernel, one crossterm (t, f) point was still selected out of the three (t, f) points. When only the autoterm (t, f) points are used, the $\hat{\mathbf{A}}_e^{\#}\mathbf{A}$ matrix becomes

$$\hat{\mathbf{A}}_e^{\#}\mathbf{A} = \begin{bmatrix} 1.00 + j0.00 & -0.02 + j0.18 \\ 0.01 + j0.14 & 1.00 + j0.00 \end{bmatrix}$$

and the global rejection level I_{perf} is reduced to -12.86 dB. The CW distributions of the separated signals are shown in Fig. 9(b).

Fig. 10 shows the CW distributions of the separated signals under the same condition as Fig. 9(a), with the utilization of the spatial averaging method. Again, it is clear that the time-frequency distributions of both cases are the same as the source signals, and $\hat{\mathbf{A}}_e^{\#}\mathbf{A}$ are exactly an identity matrix. The same results can be obtained when only the autoterm (t, f) points are used.

Fig. 11 shows the global rejection level versus the frequency difference δf between the two chirps, where the input SNR is 20dB. It is evident from this figure that the kernel method fails when the two signals have close time-frequency signatures. Using the proposed spatial averaging method outperforms the case when no spatial averaging is applied. Three important observations on the difference between the WV distribution and the CW distribution in the context of source separation are in order. First, the CW kernel effectively reduces the crossterms, particularly when δf is large. Accordingly, crossterms are not as large as the autoterms, and as such, it is unlikely for the crossterms to be selected and incorporated in the joint-diagonalization scheme. Second, when δf is large enough, the global rejection level is significantly reduced for the CW distribution, even when spatial averaging is not applied. Third, when the spatial averaging method is used, the performance at small frequency offset from the CW distribution is worse than that obtained from the WV distribution. The reason is, source separation is perturbed by the presence of noise, and the performance nevertheless is sensitive to the input SNR. When comparing the WV distribution and the CW distribution, the noise floor relative to peak values is lower in WV distribution than CW for the underlying chirp signal example.

To show the effect of the input SNR on the source separation performance, Fig.12 and Fig.13 depict the global rejection level versus the input SNR, where the frequency difference is 0.01. Increasing the SNR certainly improves the source separation performance when spatial averaging is applied. On the other hand, without spatial averaging, the source separation performance holds an almost constant high level. Such performance demonstrates that crossterms are more of a fundamental problem than noise in TFD-based source separation problems.

In the second set of simulations, ω is set to zero in equation (38), rendering the two source signals identical in terms of their instantaneous frequency characteristic. However,

one of the two source signals is amplitude modulated, caused by a nonzero positive value of α.

Fig.14 shows the global rejection level versus α, where the WV distribution is considered, and the input SNR is 20dB. It is clear that, the two signals cannot be separated without spatial averaging, but when applying spatial averaging, satisfactory performance of source separation can be achieved. For α=0.002, the proposed technique yields a global rejection level -26.72 dB.

V. Conclusions

Spatial averaging of spatial time-frequency distributions has been introduced and the role of spatial averaging in mitigating the effects of crossterms when bilinear transforms are used for signal recovery has been shown. The spatial averaging of the spatial time-frequency distributions of the data across an antenna array removes the undesired effect of crossterms between the impinging signals. These terms reside along the off-diagonal entries of the source time-frequency distribution matrix, and consequently impede the source separation performance, which is based on preassumed diagonal matrix structure. Spatial averaging amounts to forming a spatial Hermitian Toeplitze matrix using the auto- and cross-time-frequency distributions of the data over one half of the uniform linear array. This matrix is then added to the spatial matrix corresponding to the other half of the array. The desired effect of this averaging is reallocating the interaction between the source signals in the time-frequency domain from the off-diagonal to the diagonal elements of the source TFD matrix. In this respect, unlike the method proposed in [2], cross-terms, due to their potential high values, are regarded as useful components that could be used for improved performance. Spatial averaging can be applied to all members of Cohen's class of TFDs, irrespective of the employed smoothing kernel. When using a time-frequency kernel, the problem amounts to averaging in all three dimensions of time, frequency, and space.

Joint-diagonalization (JD) is applied to include multiple spatially averaged time-frequency distributions at different time-frequency points. With cross-terms moved to the diagonal entries of the TFD matrix, the prime task of the source separation based on the JD scheme is to avoid degenerate eigenvalues which are responsible for non-uniqueness solution of the problem.

Simulation examples were presented to illustrate the effectiveness of the new approach. The two performance measures used were the global rejection level and the values of the off-diagonal elements of the product of the mixing matrix and the Pseudo inverse of its estimate. Two sources and five sensors were considered. The source signals were chirp signals with the same sweeping frequency, but their corresponding constant frequencies and amplitudes were offset by different values. Both Wigner-Ville and Choi-Williams distributions were considered. It was shown that the spatial averaging method significantly improves the performance measures over the non-spatially averaging method, specifically when the two signals have close time-frequency signatures.

Without spatial-averaging, performance is very sensitive to whether only auto-term or cross-term points or their mix are incorporated in the source separation procedure. With spatial averaging, this is no longer a concern, since both terms appear along the diagonal. It is also shown that the Choi-Williams distribution provides better results than the Wigner-Ville distribution when no spatial averaging is applied, since it lowers the likelihood of selecting crossterm points. With spatial averaging, the issue becomes merely SNR, and in this respect, the Wigner-Ville distribution, due to its high peak values, yields better performance than the Choi-Williams distribution. Therefore, the time-frequency smoothing becomes unnecessary whenever spatial array averaging is possible.

References

[1] A. Belouchrani and M. Amin, "Source separation based on the diagonalization of a combined set of spatial time-frequency distribution matrices," in *Proc. ICASSP'97*, Germany, April 1997.

[2] ——, "Blind source separation based on time-frequency signal representation," *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. 2888-2898, Nov. 1998.

[3] ——, "Blind source separation using joint signal representations," in *Proc. SPIE: Advanced Algorithms and Architectures for Signal Processing*, San Diego, CA, Aug. 1997.

[4] M. G. Amin and A. Belouchrani, "Blind source separation using the spatial ambiguity functions," in *Proc. IEEE Int. Symp. on Time-Frequency and Time-Scale Analysis*, Pittsburgh, PA, pp. 413–416, Oct. 1998.

[5] A. Belouchrani and M. Amin, "Time-frequency MUSIC," *IEEE Signal Processing Letters*, vol. 6, no. 5, pp. 109–110, May 1999.

[6] S. U. Pillai, *Array Signal Processing*, Springer-Verlog, 1989.

[7] L. Cohen, *Time-frequency Analysis*, Prentice Hall, 1995.

[8] L. Tong, Y. Inouye, and R-W. Liu, "Waveform-preserving blind estimation of multiple independent sources," *IEEE Trans. Signal Processing*, vol, 41, no. 7, pp. 2461–2470, July 1993.

[9] A. Belouchrani, K. A. Meraim, H.-F. Cardoso, and E. Muylines, "A blind source separation techniques using second order statistics," *IEEE Trans. Signal Processing*, vol. 45, no. 2, pp. 434–444, Feb. 1997.

[10] H. I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 6, pp. 862–871, June 1989.

[11] P. Stoica and K. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 7, pp. 1132–1143, July 1990.

[12] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 5, pp. 814–824, May 1990.

[13] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. 34, no. 3, pp. 276–280, March 1986.

[14] A. J. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. ICASSP'83*, Boston, MA, pp. 336–339, 1983.

[15] B. D. Rao and K. V. S. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 12, pp. 1939–1949, Dec. 1989.

[16] R. Roy and T. Kailath, "ESPRIT - estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 984–995, July 1989.

[17] Y. Zhang, W. Mu, and M. G. Amin, "Subspace analysis of spatial time-frequency distribution matrices," submitted to *IEEE Trans. Signal Processing.*

[18] Y. Zhang, W. Mu, and M. G. Amin, "Time-frequency maximum likelihood methods for direction finding," to appear at *Journal of the Franklin Institute.*

[19] Y. Zhang and M. G. Amin, "Blind separation of sources based on their time-frequency signatures," in *Proc. ICASSP'00*, Istanbul, Turkey, June 2000.

[20] J. E. Evans, J. R. Johnson, and D. F. Sun, "High resolution angular spectrum estimation techniques for terrain scattering analysis and angle of arrival estimation," in *Proc. 1st ASSP Workshop on Spectral Estimation*, Hamilton, Canada, Aug. 1981.

[21] T. J. Shan and T. Kailath, "On spatial smoothing for DOA estimation of coherent sources," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 4, pp. 806–811, 1985.

[22] R. T. Williams, S. Prasad, A. K. Mahalanabis, and L. H. Sibul, "An improved spatial smoothing technique for bearing estimation in multipath environment," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 4, pp. 425–432, 1988.

[23] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 1, pp. 8–15, Jan. 1989.

[24] S. U. Pillai, Y. Barness, and F. Haber, "A new approach to array geometry for improved spatial spectrum estimation," *Proc. IEEE*, vol. 73, pp. 1522–1524, Oct. 1985.

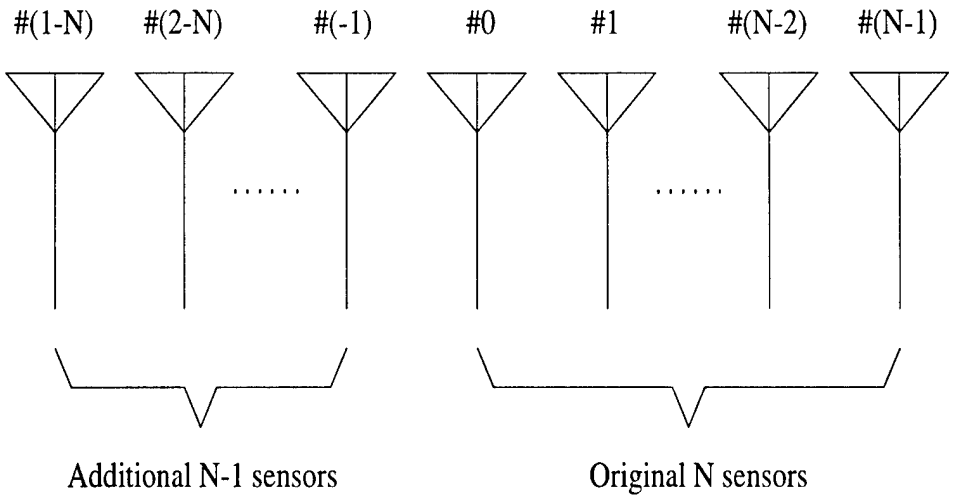


Fig. 1 Array configuration for spatial averaging

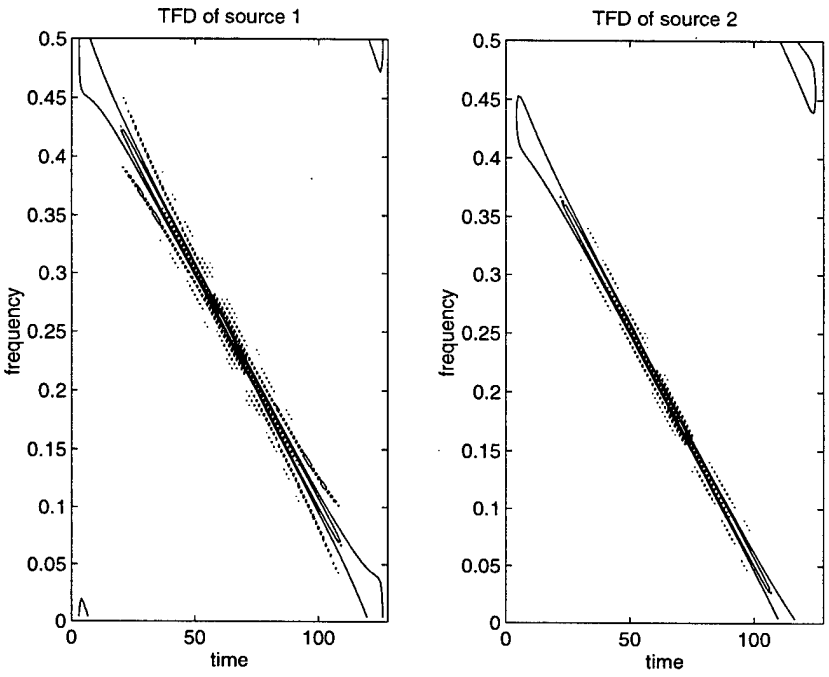


Fig. 2 TFD of the source signals (Wigner-Ville distribution)

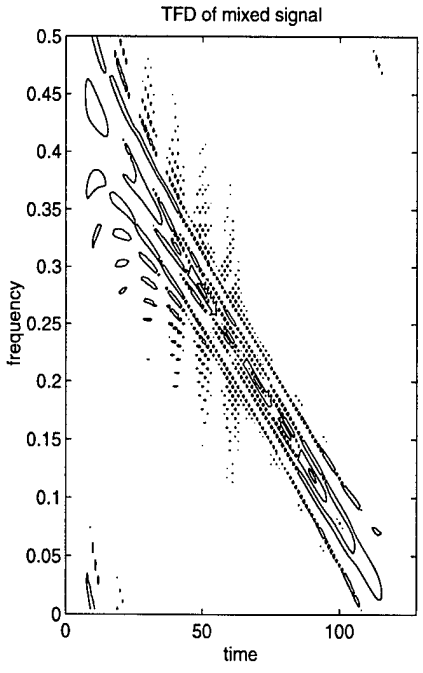
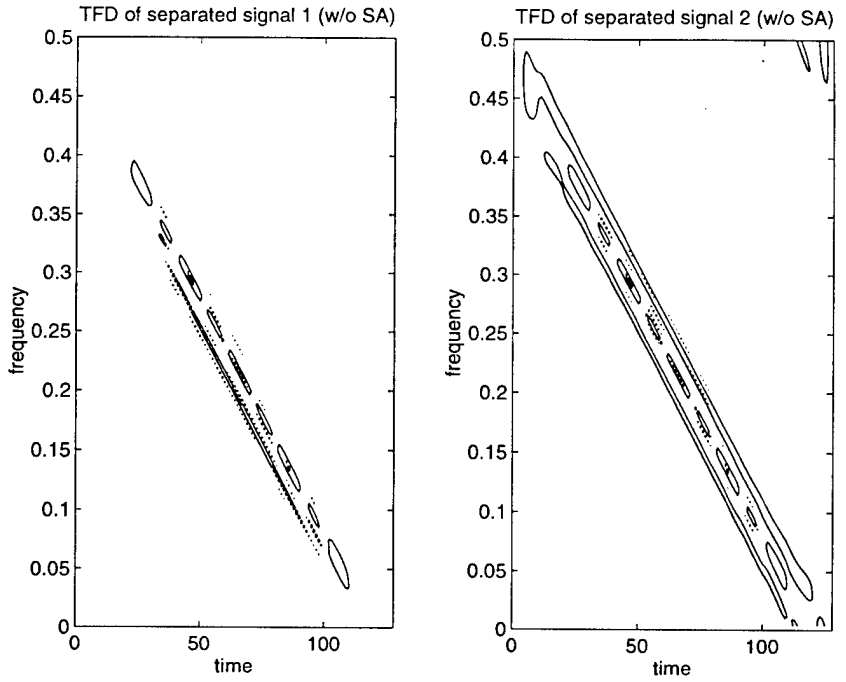
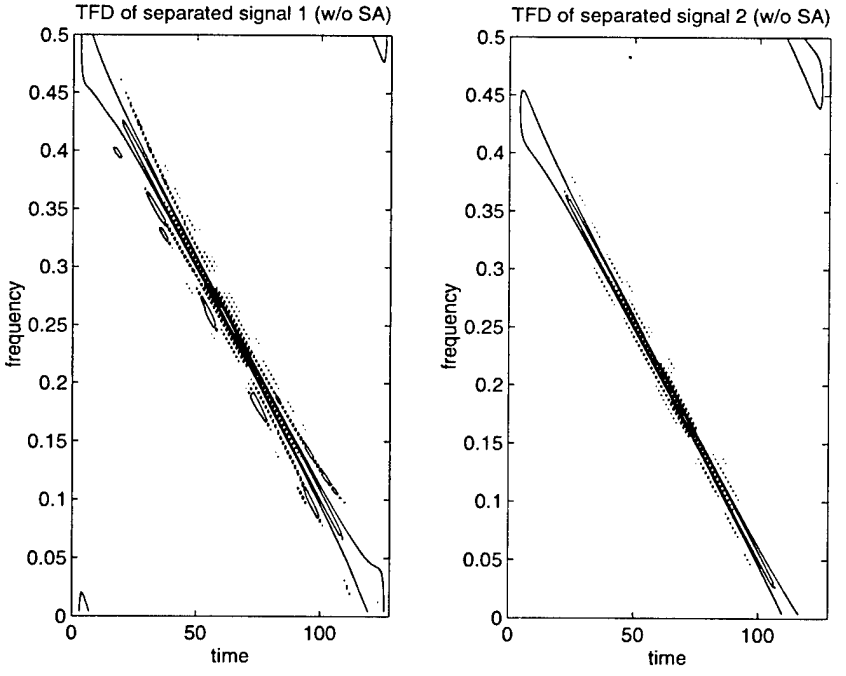


Fig. 3 TFD of the mixed signal (Wigner-Ville distribution)



(a) Using peak time-frequency points



(b) Using autoterm points

Fig. 4 TFD of the separated signals without spatial averaging (Wigner-Ville distribution)

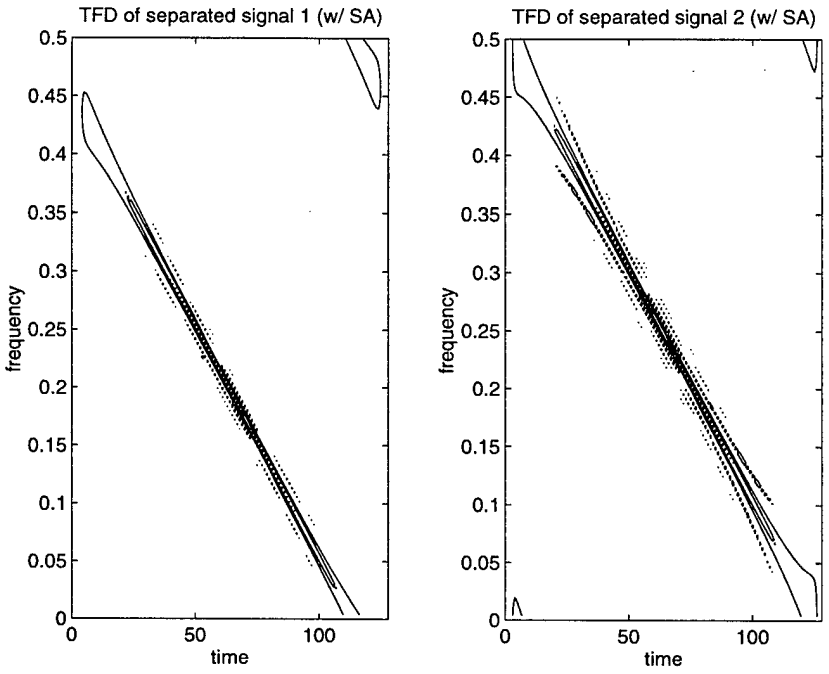


Fig. 5 TFD of the separated signals with spatial averaging (Wigner-Ville distribution)

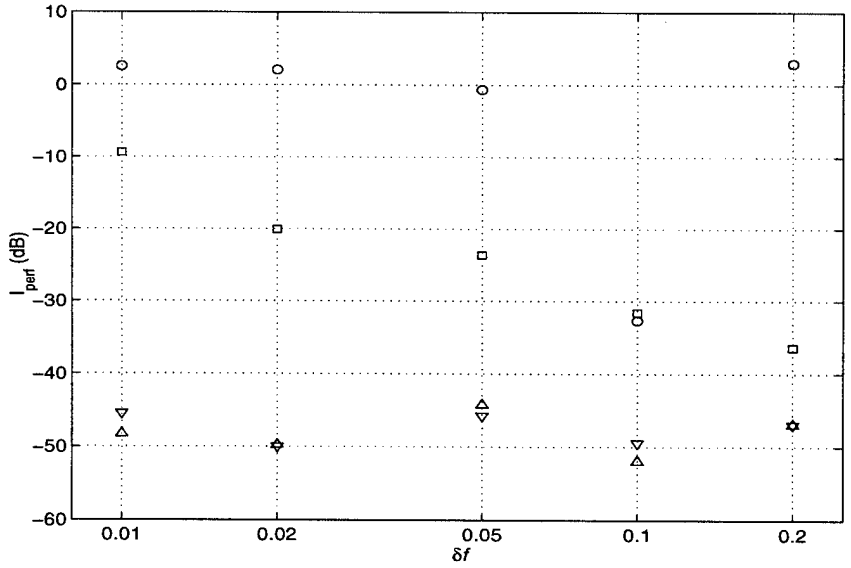


Fig. 6 Global rejection level versus frequency difference δf (Wigner-Ville distribution)
(input SNR = 20 dB, o: without spatial averaging, $\triangle$: with spatial averaging,
$\square$: without spatial averaging using autoterm points,
$\triangledown$: with spatial averaging using autoterm points)

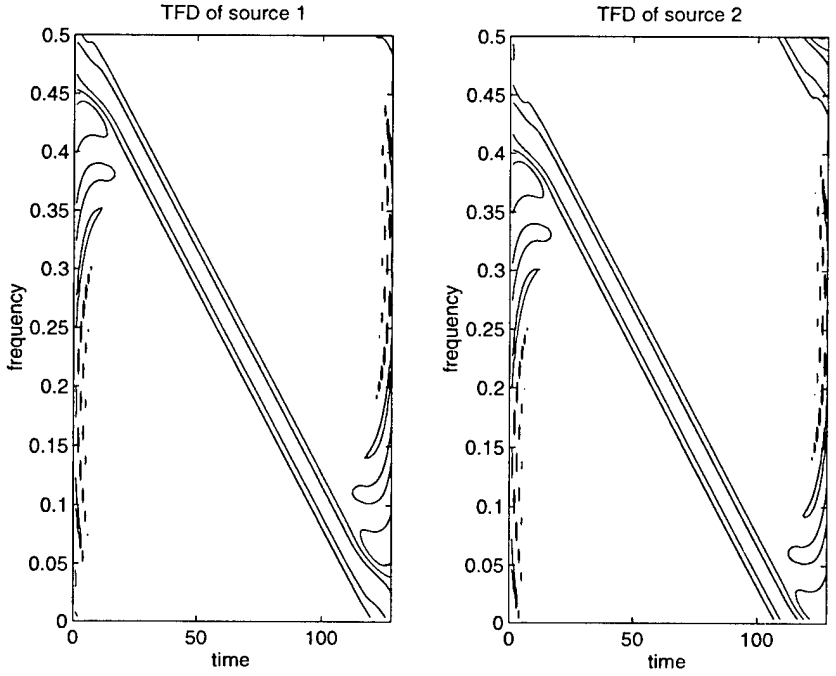


Fig. 7 TFD of the source signals (Choi-Williams distribution)

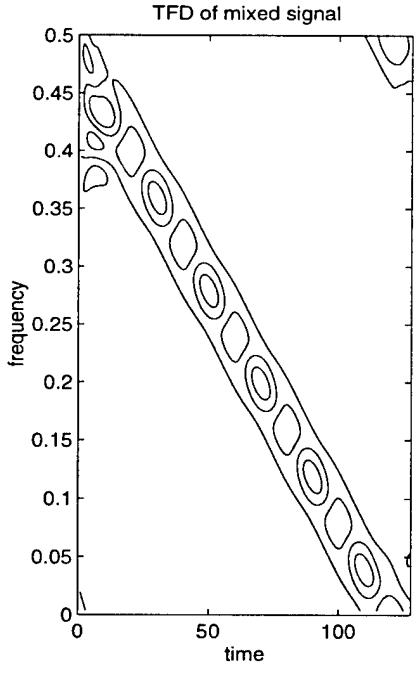
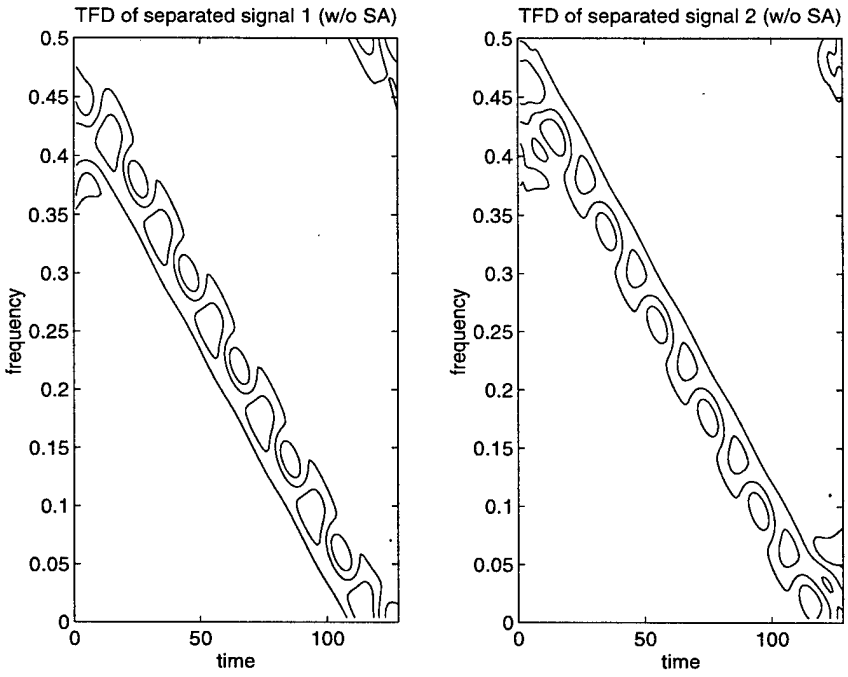
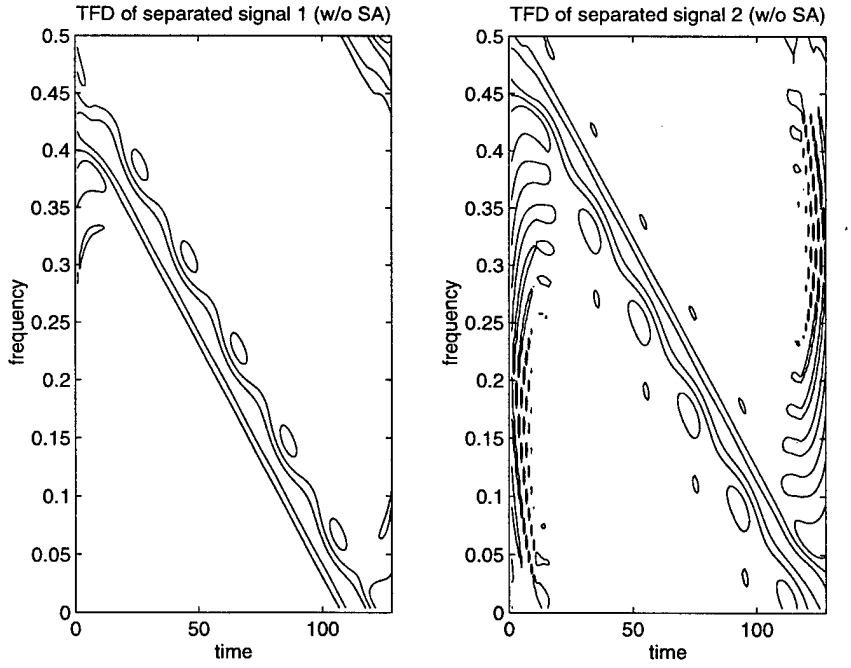


Fig. 8 TFD of the mixed signal (Choi-Williams distribution)



(a) Using peak time-frequency points



(b) Using autoterm points

Fig. 9 TFD of the separated signals without spatial averaging (Choi-Williams distribution)

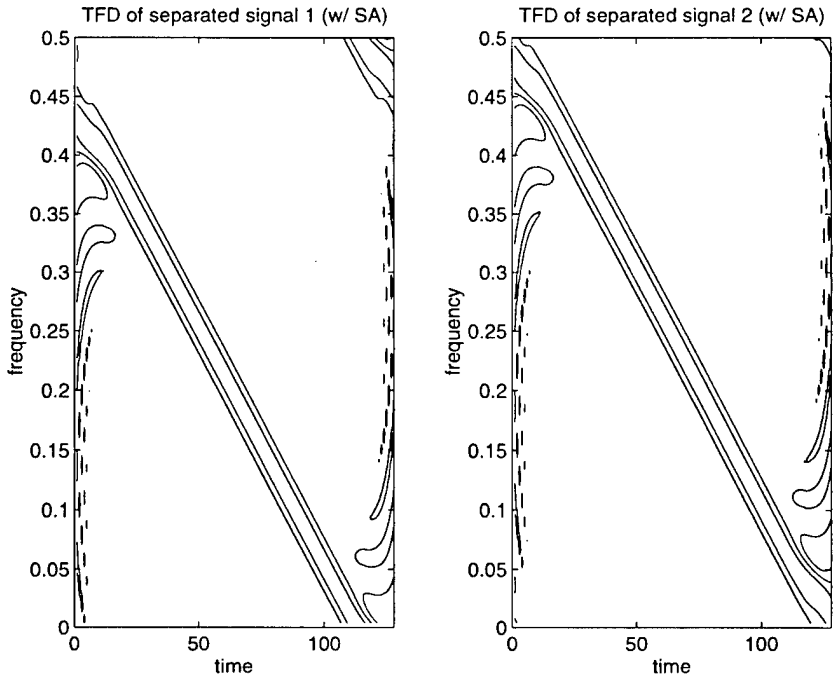


Fig. 10 TFD of the separated signals with spatial averaging (Choi-Williams distribution)

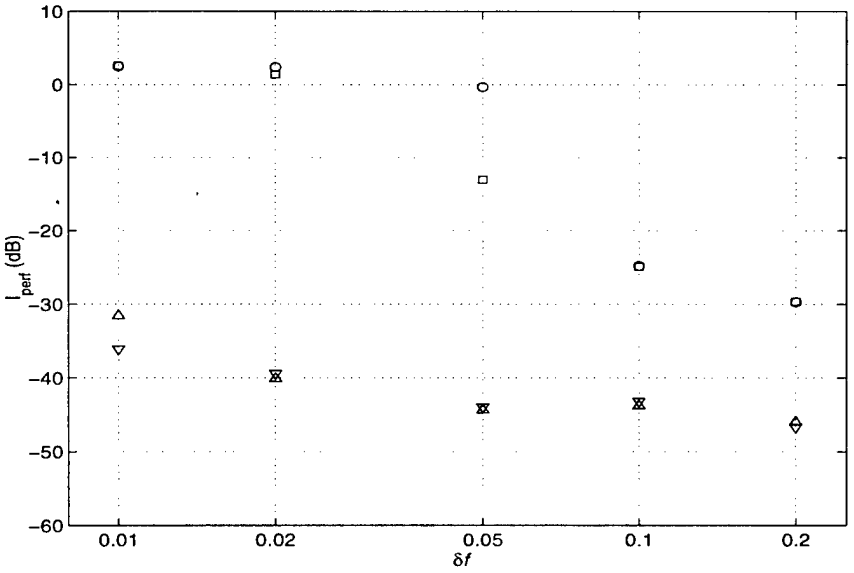


Fig. 11 Global rejection level versus frequency difference δf (Choi-Williams distribution)

(input SNR = 20 dB, o: without spatial averaging, △: with spatial averaging,

□: without spatial averaging using autoterm points,

▽: with spatial averaging using autoterm points)

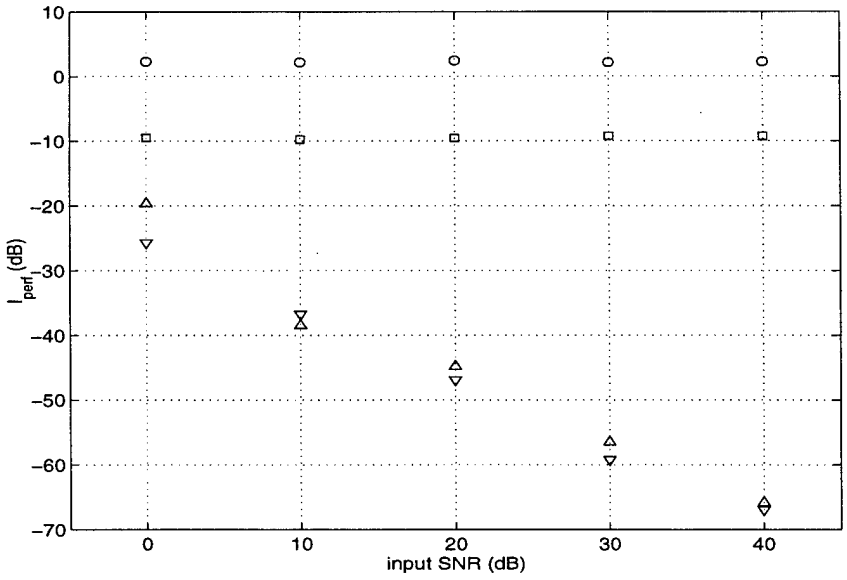


Fig. 12 Global rejection level versus input SNR (Wigner-Ville distribution)
(δf=0.01, o: without spatial averaging, $\triangle$: with spatial averaging,
$\square$: without spatial averaging using autoterm points,
$\bigtriangledown$: with spatial averaging using autoterm points)

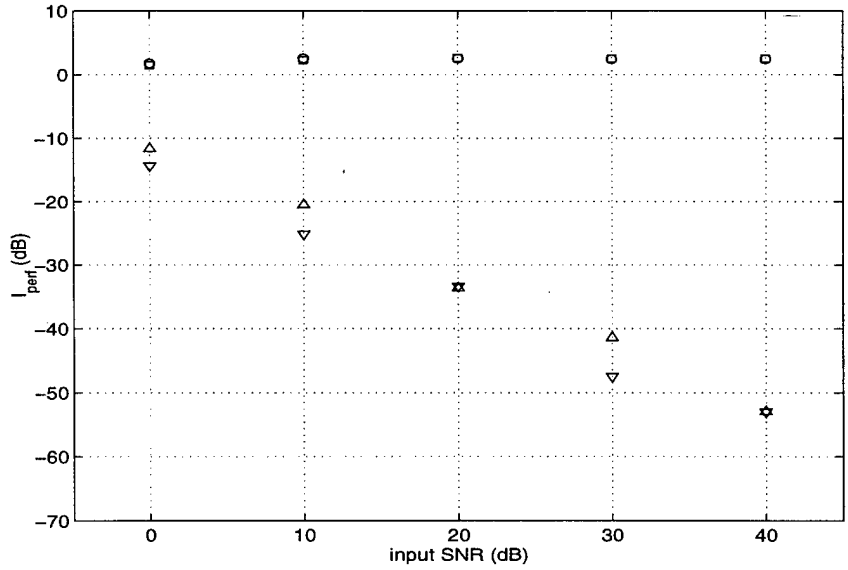


Fig. 13 Global rejection level versus input SNR (Choi-Williams distribution)
(δf=0.01, o: without spatial averaging, $\triangle$: with spatial averaging,
$\square$: without spatial averaging using autoterm points,
$\bigtriangledown$: with spatial averaging using autoterm points)

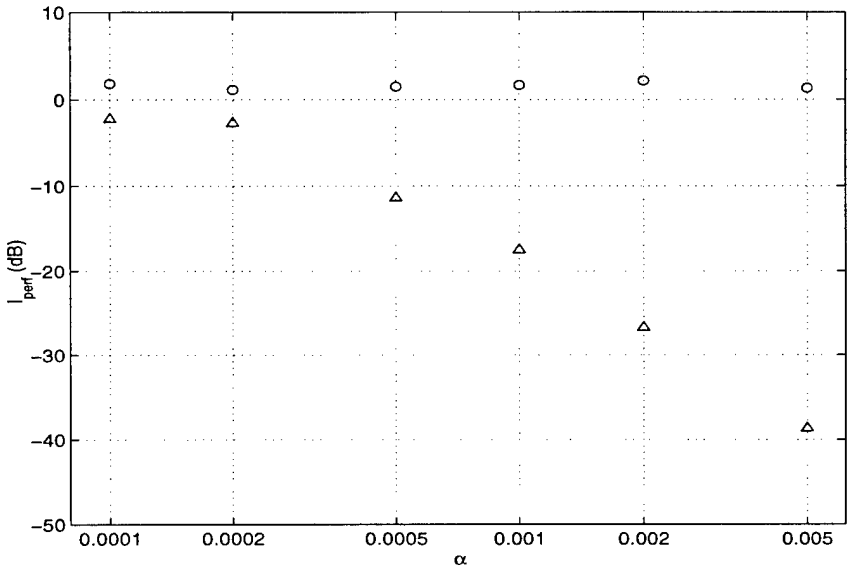


Fig. 14 Global rejection level versus α (Wigner-Ville distribution)
(input SNR = 20 dB, o: without spatial averaging, $\triangle$: with spatial averaging)

Bilinear Signal Synthesis in Array Processing

Weifeng Mu, Moeness G. Amin, and Yimin Zhang

ECE Department, Villanova University, Villanova, PA 19085

Abstract

Multiple source signals impinging on an antenna array can be separated by time-frequency synthesis techniques. Averaging of the time-frequency distributions of the data across the array permits the spatial signatures of sources to play a fundamental role in improving the synthesis performance. Array averaging introduces a weighing function in the time-frequency (t-f) domain that decreases the noise levels, reduces the interactions of the source signals, and mitigates the crossterms. This is achieved independent of the temporal characteristics of the source signals and without causing any smearing of the signal terms. The weighing function may take non-integer values, which are determined by the communication channel, the source positions and their angular separations. Unlike the recently devised blind source separation methods using spatial time-frequency distributions, the proposed method does not require whitening or retrieval of the source directional matrix. The paper evaluates the proposed method in terms of performance and computations relative to the existing source separation techniques based on quadratic t-f distributions.

Keywords

Signal synthesis, time-frequency distribution, array processing, source separation.

This work is supported by the Office of Naval Research under Grant no. N00014-98-1-0176 and by the Air Force Research Laboratory under Grant no. F30602-00-1-0515.

I. Introduction

Time-frequency distributions have been shown to be very useful for nonstationary signal analysis and synthesis [1], [2], [3], [4], [5]. While time-frequency distributions (TFDs) have been sought out and successfully used in the areas of speech, biomedicine, automotive industry, and machine monitoring, their applications to sensor and spatial signal processing have not been sufficiently investigated. By properly incorporating the spatial dimension into time-frequency signal representations, the bilinear and higher order forms of TFD can be a powerful tool for high resolution angle of arrival estimation and recovery of the source waveforms impinging on a multi-sensor receiver, specifically those of nonstationary temporal characteristics.

Synthesizing the signal from bilinear distributions of the data at a single antenna receiver is often impeded by the presence of high levels of noise and crossterms. These undesired terms not only obscure the true signal power localization in the time-frequency (t-f) domain, but also reduce the synthesized signal quality. Signal synthesis using TFDs can be improved using an antenna array receiver. The availability of the source signals at different array elements allows the implementation of t-f synthesis techniques that utilize the source spatial signatures for crossterm reduction and noise mitigation. In this paper, we introduce a new approach for signal synthesis in antenna arrays which utilizes the spatial separation of the sources as well as the sources' time-frequency characteristics. In effect, we perform source separation, or signal recovery, based on the difference in both the time-frequency and spatial signatures of the signal arrivals. The signals impinging on the multi-antenna receiver are assumed to be localizable in the time-frequency domain, e.g., FM and polynomial phase signals.

Unlike the proposed technique, the existing array signal processing techniques for nonstationary source separation using bilinear distributions require the construction of spatial time frequency distribution (STFD) matrices from the data snap shots. The elements of

this matrix represent the auto- and cross- time frequency distributions of the data across the array. It was shown in [12], [13], [14] that the formula relating the TFD matrix of the sensor data to that of the sources is identical to the relationship between the data covariance matrix and the source correlation matrix. Blind source separation (BSS) can therefore be performed using the source t-f signatures, instead of their correlation functions. The former is more suitable for nonstationary signal environments. The BSS based on TFDs method introduced in [12] first estimates the array, or the spatial signature, matrix from the STFD using joint diagonalization. Then, it proceeds to use this matrix estimate to undo the mixing at the array and recover the source signals.

The main difficulty of the above approach, however, is the need to construct the STFD matrices from auto-term points. Selections of cross-terms violate the diagonal structure of the source TFD matrix – a necessary condition for most blind source separation methods. Even if successfully selected, the autoterm region is often contaminated by high level of noise and intruded upon by the crossterms through the energy in their mainlobes and/or sidelobes. The key feature of the proposed technique is the utilization of the sources' spatial structures to enhance their time-frequency signatures in the t-f domain. Bilinear signal synthesis methods [6], [7], [8] can then be applied to the enhanced source t-f features to recover the signal waveform and its temporal characteristics. By averaging the time-frequency distributions of the data across the array, we permit the source spatial signatures (SS) to play a fundamental role in reducing noise and crossterm contamination of the true signal t-f power concentration, leading to improved synthesis performance. It is shown that the performance is determined by the inner product of the source array vectors and improves for weakly correlated and orthogonal source spatial signatures. In the case of Gaussian channel and omni-directional uniform linear arrays, spatial averaging produces a *sinc* weighting function whose maximum value, normalized to one, is assigned to all source autoterms, whereas its fractional values are assigned to the source crossterms, and thereby mitigating their effects. It is shown that the extent of crossterm reduction is a

function of the spatial frequency separation of the sources and does not rely on the source specific locations or their time-frequency characteristics. However, sources positioned near the broadside will generally exhibit lower interactions in the t-f domain than those at the endfire.

Unlike source separation techniques based on STFD, the proposed approach does not require whitening or retrieval of the source directional matrix, thereby, simplifies the signal recovery process. Further, as a result of the averaging process in the proposed approach, a weighting function in the time-frequency domain is constructed which decreases the noise levels, reduces the interactions of the source signals, and mitigates the cross-terms. This is achieved independent of the temporal characteristics of the source signals and without causing any smearing of the signal auto-terms.

The paper is organized as follows. The signal model is presented in Section II, and the proposed array averaging technique is also formulated. The effect of source angular separation on cross-term reduction is cast in Section II using the implicit beamforming properties of spatial averaging. Section II also addresses the equivalent t-f mask introduced by the proposed technique. The complete synthesis procedure is devised in Section III, where the signals are synthesized from the array averaged modified WVD. The modified WVD [9] is used to avoid the need for extracting the odd-indexed and even-indexed vectors separately via eigen-analysis. Numerical simulations illustrating the performance of the proposed method are given in Section IV.

II. Problem Formulation

A. Signal Model

Assume L source signals incident on an M-sensor array. The propagation delay between antenna elements is assumed to be small relative to the inverse of the transmission bandwidth, so that the received signals are identical to within a complex constant. The data

received across the array is given by the narrowband model

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad t = 1, \ldots, N \tag{1}$$

where $\mathbf{x}(t) = [\mathrm{x}_1(t), \ldots, \mathrm{x}_M(t)]^T$ and $\mathbf{s}(t) = [\mathrm{s}_1(t), \ldots, \mathrm{s}_L(t)]^T$ are the $M \times 1$ data snapshot vector and the $L \times 1$ source signal vector at time instant t, respectively. The superscript ^{T} denotes the vector/matrix transpose. The $M \times 1$ vector $\mathbf{n}$(t) is the noise vector whose elements are modeled as stationary, spatially and temporally white complex Gaussian processes with zero mean and variance of σ^2, i.e.,

$$E\left[\mathbf{n}(t+\tau)\mathbf{n}^H(t)\right] = \sigma^2\delta(\tau)\mathbf{I} \tag{2}$$

where the superscript ^{H} denotes transpose conjugation. Moreover, $\delta(\tau)$ is the kronecker delta and $\mathbf{I}$ denotes the identity matrix, and $\mathbf{A}$ denotes the $M \times L$ mixing matrix,

$$\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_L], \tag{3}$$

The columns of matrix $\mathbf{A}$ are the source spatial signatures (SSs), and are given by

$$\mathbf{a}_i = [\mathrm{a}_{i1}, \ldots, \mathrm{a}_{iM}]^T \tag{4}$$

where a_{ij} is the jth component of the ith SS, $\mathbf{a}_i$. Matrix $\mathbf{A}$ serves as the transfer function between the source signals $\mathbf{s}$(t) and the data $\mathbf{x}$(t). Furthermore, we assume that matrix $\mathbf{A}$ is of full column rank, which implies that the SSs associated with the L sources are linearly independent. To simplify the discussion, we exchange any possible scalar factor embedded in $\mathbf{a}_i$ to the source signal and assume that $||\mathbf{a}_i||_2 = M$. It is clear that this exchange does not affect the data observed from the antenna array.

It is evident that when $L > 1$, equation (1) represents a multi-component scenario due to the mixture of the signals at each sensor. Therefore, a quadratic TFD at the individual sensors would contain not only the autoterms of all source signals, but also the interactions of the source signals, causing undesirable crossterms.

For the purpose of subsequent derivation, we first expand equation (1) using definitions (3) and express the received noise-free data vector

$$\mathbf{y}(t) = \sum_{i=1}^{L} \mathbf{a}_i \mathrm{s}_i(t). \tag{5}$$

Specifically, the data received at sensor k ($k = 1, 2, \ldots, M$) is given by

$$\mathrm{y}_k(t) = \sum_{i=1}^{L} \mathrm{a}_{ik} \mathrm{s}_i(t). \tag{6}$$

B. The Array Averaged WVD

The discrete form of WVD of the signal $\mathrm{y}(t)$ is given by [4]

$$\mathrm{W}_{\mathrm{yy}}(t, f) = \sum_{l=-\infty}^{\infty} \mathrm{y}(t+l)\mathrm{y}^*(t-l)e^{-j4\pi fl}, \tag{7}$$

where * denotes complex conjugation, and t and f represent the time index and the frequency index, respectively. Equation (7) is often referred to as the auto WVD of the signal $\mathrm{y}(t)$. Similarly, the cross WVD of any two signals $\mathrm{y}_1(t)$ and $\mathrm{y}_2(t)$ is defined as

$$\mathrm{W}_{\mathrm{y_1y_2}}(t, f) = \sum_{l=-\infty}^{\infty} \mathrm{y}_1(t+l)\mathrm{y}_2^*(t-l)e^{-j4\pi fl}. \tag{8}$$

Substituting (6) into (7), we can express the WVD of the signal at the kth sensor $\mathrm{y}_k(t)$ as

$$\mathrm{W}_{\mathrm{y}_k\mathrm{y}_k}(t, f) = \sum_{i=1}^{L}\sum_{j=1}^{L} \mathrm{a}_{ik}\mathrm{a}_{jk}^* \sum_{l=-\infty}^{\infty} \mathrm{s}_i(t+l)\mathrm{s}_j^*(t-l)e^{-j4\pi fl} = \sum_{i=1}^{L}\sum_{j=1}^{L} \mathrm{a}_{ik}\mathrm{a}_{jk}^* \mathrm{W}_{\mathrm{s}_i\mathrm{s}_j}(t, f), \tag{9}$$

where $\mathrm{W}_{\mathrm{y}_k\mathrm{y}_k}(t, f)$ will herein be referred to as the auto-sensor WVD of $\mathrm{y}_k(t)$. $\mathrm{W}_{\mathrm{s}_i\mathrm{s}_j}(t, f)$ corresponds to the auto-source or cross-source WVD, depending on whether $i = j$, or $i \neq j$. It is important to note that there are two types of crossterms in the underlying problem. The first type are the crossterms that are present in the auto-source WVD. These terms are the results of the interactions between the components of the same source signal, which is the case when the source signal itself is of multi-components. Without loss of generality, we assume mono-component sources, and as such the auto-source WVD has no cross-terms. The other type of crossterms are found in the cross-source WVD and

generated from the interactions between two signal components belonging to two different sources. A variant of the two is the crossterms in the cross-sensor WVD, which results from the interactions of the signals from different array elements. It is generally the collection of the above crossterms and the source auto-terms. We note that the cross-sensor WVD does not play a role in the analysis presented in this paper.

Averaging the auto-sensor WVDs over the array yields

$$\begin{aligned}\bar{\mathrm{W}}(t,f) &= \frac{1}{M}\sum_{k=1}^{M}\mathrm{W}_{\mathrm{y}_k\mathrm{y}_k}(t,f) = \sum_{i=1}^{L}\sum_{j=1}^{L}\left(\frac{1}{M}\sum_{k=1}^{M}\mathrm{a}_{ik}\mathrm{a}_{jk}^{*}\right)\mathrm{W}_{\mathrm{s}_i\mathrm{s}_j}(t,f) \\ &= \sum_{i=1}^{L}\sum_{j=1}^{L}\left(\frac{1}{M}\mathbf{a}_j^H\mathbf{a}_i\right)\mathrm{W}_{\mathrm{s}_i\mathrm{s}_j}(t,f). \qquad (10)\end{aligned}$$

In equation (10), $\mathbf{a}_j^H\mathbf{a}_i$ is the inner product of the SSs $\mathbf{a}_i$ and $\mathbf{a}_j$. For $i = j$, $\mathbf{a}_i^H\mathbf{a}_i = ||\mathbf{a}_i||_2 = M$. Define the spatial correlation coefficient

$$\beta_{ij} = \frac{1}{M}\mathbf{a}_j^H\mathbf{a}_i, \qquad (11)$$

Equation (10) can be then rewritten as

$$\bar{\mathrm{W}}(t,f) = \sum_{i=1}^{L}\sum_{j=1}^{L}\beta_{ij}\mathrm{W}_{\mathrm{s}_i\mathrm{s}_j}(t,f). \qquad (12)$$

The above equation shows that $\bar{\mathrm{W}}(t,f)$ is a linear combination of the auto-source and cross-source WVDs of all signal arrivals. To obtain a general and compact form for $\bar{\mathrm{W}}(t,f)$, we define the source WVD matrix that enters $\mathrm{W}_{\mathrm{s}_i\mathrm{s}_j}(t,f)$ as its (i,j)th element,

$$\mathbf{W}_{\mathbf{ss}}(t,f) = \left[\mathrm{W}_{\mathrm{s}_i\mathrm{s}_j}(t,f)\right], \quad i,j = 1,\ldots,L \qquad (13)$$

and

$$\mathbf{\Gamma} = \frac{1}{M}\mathbf{A}^H\mathbf{A} = [\beta_{ij}], \quad i,j = 1,\ldots,L \qquad (14)$$

Accordingly, the averaged WVD $\bar{\mathrm{W}}(t,f)$ could be simplified to

$$\bar{\mathrm{W}}(t,f) = \mathbf{u}^H\left[\mathbf{W}_{\mathbf{ss}}(t,f)\odot\mathbf{\Gamma}\right]\mathbf{u}, \qquad (15)$$

where $\odot$ denotes the Hadamard product, or the matrix element-by-element product, and $\mathbf{u}$ is an $L \times 1$ vector of unit values, i.e., $\mathbf{u} = [1, \ldots, 1]^T$. Equation (15) is valid for every (t, f) point and elucidates the averaging of the WVD across the array. It includes all the signal autoterms and crossterms that naturally appear in a typical multi-component WVD. However, in (15), these autoterms and crossterms are weighted by constant values, which are the spatial correlation coefficients that have resulted from the inner product between the sources' SSs, exhibited in the elements of matrix $\mathbf{\Gamma}$. It is important to note that by the virtue of the inner product, the source directional information carried by its respective SS is lost in $\bar{\mathrm{W}}(t, f)$.

The diagonal elements of the matrix $\mathbf{W}_{\mathbf{ss}}(t, f) \odot \mathbf{\Gamma}$ constitute all the autoterms of the L source signals, whereas the off-diagonal elements are their respective cross-terms. It is straightforward to show that for the ith and the jth sources,

$$|\beta_{ij}| \leq 1, i \neq j \quad \text{and} \quad \beta_{ij} = 1, i = j, \tag{16}$$

indicating that the constant coefficients in (15) for the auto-source WVDs are always greater than, or at least equal to, those for the cross-source WVDs. For a large array or widely separated sources, $|\beta_{ij}| \ll 1$, leading to significant suppression of the crossterms. This property is utilized by the array averaging process and is shown to improve the signal synthesis performance.

An interesting case arises when all SSs are orthogonal, i.e., $\beta_{ij} = 0$ for any $i \neq j$. In this case, $\mathbf{\Gamma}$ becomes an identity matrix and yields

$$\bar{\mathrm{W}}(t, f) = \sum_{k=1}^{L} \mathrm{W}_{\mathrm{s}_k \mathrm{s}_k}(t, f). \tag{17}$$

It should be noted that $\bar{\mathrm{W}}(t, f)$ in (17) is solely the summation of the source signal autoterms. The above equation highlights the fact that all source signal crossterms are entirely eliminated from $\bar{\mathrm{W}}(t, f)$ and only the autoterms are maintained, which is most desirable from the synthesis perspective.

C. Beamforming Effect

In order to establish quantified analysis on the suppression effect of crossterms based on the proposed array averaging technique, we consider the special case of a Gaussian channel and a uniform linear array(ULA). With no signal scattering, the SS displays the structure

$$\mathbf{a}_i = \left[1, e^{j\omega_i}, \ldots, e^{j(M-1)\omega_i}\right]^T. \tag{18}$$

The spatial frequency of the ith source ω_i is given by

$$\omega_i = \frac{2\pi d}{\lambda} \sin(\theta_i), \tag{19}$$

where d is the inter-element spacing, λ is the wavelength, and θ_i is the angle of arrival (AOA). From (11) we obtain

$$|\beta_{ij}| = \left| \frac{1 - e^{jM\Delta\omega}}{M\left(1 - e^{j\Delta\omega}\right)} \right| = \frac{\sin(M\Delta\omega/2)}{M \cdot \sin(\Delta\omega/2)} \tag{20}$$

where $\Delta\omega = \omega_j - \omega_i$ denotes the difference between the two spatial frequencies ω_j and ω_i. Equation (20) is the well-known *array factor* for an M-element linear array [15]. The spatial pattern represented by (20) reaches its maximum value within the mainlobe at $\Delta\omega = 0$. The pattern has secondary maxima in the side lobes. The largest of those maxima occurs within the first sidelobe and is asymptotically (for large M) 13dB down from the unit value. In this regard, if the difference in the spatial frequencies of adjacent sources are greater than $\Delta\Omega = \frac{2\pi}{M}$, which is half of the mainlobe width, the suppression of crossterms could always be guaranteed by at least 13dB for a large value of M. From equation (20), we have

$$\lim_{M\to\infty} |\beta_{ij}| = 0, \quad i \neq j. \tag{21}$$

Using this result, we could further rewrite equation (12) as

$$\lim_{M\to\infty} \bar{\mathrm{W}}(t,f) = \sum_{k=1}^{L} \mathrm{W}_{\mathrm{s}_k\mathrm{s}_k}(t,f), \tag{22}$$

which is the asymptotical form of the orthogonal scenario described in (17). The importance of equation (22) lies in the fact that by utilizing the array averaging approach, crossterm could be suppressed to any extent if there are sufficient number of array elements. In other words, the orthogonality in SSs could be always approached by increasing the array manifold. As such, in the underlying problem, the array size is cast as an important parameter influencing the performance of crossterm suppression through array averaging.

Since M and d appear as a product in (20), then as the number of the sensor M increases, the sources could be more closely spaced without impeding crossterm suppression. It is important to note that because of the nonlinearity between $\Delta\omega$ and $\Delta\theta$, the condition on angular separation $\Delta\theta$ for the same level of cross-term suppression is more relaxed when the sources are near the broadside and more rigid when they are placed at the endfire.

The condition for the orthogonal structure of all SSs in a ULA, and subsequently full crossterm suppression, could be determined by simply setting equation (20) equal to 0. Consequently, we obtain,

$$\Delta\omega = \pm m\frac{2\pi}{M} = \pm m\Delta\Omega, \quad m = 1, 2, \ldots, \tag{23}$$

and equation (17) follows. Conversely, the worst performance corresponds to the case in which the sources have the same AOA, i.e., $\Delta\omega = 0$, or they are closely spaced, i.e., $|\Delta\omega| \ll \frac{2\pi}{M}$. In this case, $|\beta_{ij}| \approx 1$, and the crossterms would not encounter any significant changes as a result of array averaging. In general, if all the sources impinging on the array are closely spaced, the source signal crossterms will not suffer any substantial reduction, and the averaged WVD becomes almost no different from the WVD computed from the output of a single antenna receiver.

D. Equivalent Time-Frequency Mask

The incorporation of the source SSs into WVD results in multiplying the original WVD with an appropriate mask function in the t-f domain. However, unlike the conventional mask functions that typically, but not necessarily, assigns 1 for desired t-f regions and 0 elsewhere, the mask produced by the source spatial structure may assume any values, which are dependent on both the communication channels and source spatial locations. These mask values are high over autoterm regions and small over crossterm regions, regardless of their specific locations in the t-f domain.

It is noteworthy that since the crossterm suppression is controlled by the inner product of the source SSs and is not dependent on the source temporal characteristics and signal frequency contents, the evolved time-frequency mask in the underlying problem only reduces the crossterms that are produced from the interaction of the signals of different sources. That is, the array averaging process of the sensors' WVDs does not reduce the crossterms of the signal components belonging to the same source. These crossterms are, in essence, highlighted by the same coefficient that multiplies the respective source autoterms, and their retainment is important when synthesizing multi-component signals. In other words, unlike reduced interference distributions (RIDs) [4], [10] in which appropriate kernels are applied for smoothing all crossterms, the proposed synthesis method using array processing applies selective mitigation of crossterms, as it identifies and eliminates the "undesired" crossterms over any t-f regions, even if they are overlapped by the source autoterms, which is the case shown in the simulation section. Another advantage of the proposed method over the RIDs lies in the fact that, the array averaging technique does not produce any smearing effect. That is, the averaging process, apart from scalar multiplication, does not alter the shapes of the signal autoterms in the t-f domain.

The averaged WVD $\bar{\mathrm{W}}(t,f)$ is not a valid WVD, as there rarely exists a signal waveform that has the same WVD as $\bar{\mathrm{W}}(t,f)$. In the sense of signal synthesis or blind source

separation, the criterion of minimum square error(MSE) is often applied to obtain the signal $y_{opt}(t)$ with WVD that best approximates the modified WVD, i.e.,

$$y_{opt}(t) = \arg \min_{y} \sum_{t} \sum_{f} \left| W_y(t,f) - \bar{W}(t,f) \right|^2 . \tag{24}$$

If desired, one may combine kernel smoothing and spatial smoothing. This is achieved by replacing the WVD in (13-17) with another member of Cohen's class of TFDs. It is evident from above equations, however, that the extent to which the crossterms are mitigated via spatial averaging is kernel blind. It depends exclusively on the expression of $\mathbf{\Gamma}$, which is determined by the spatial information of the source signals and is independent of $\mathbf{W_{ss}}$, the specific t-f expression used. The integration of both spatial and time-frequency averaging can result in significant crossterm suppression that cannot be achieved by each type of averaging applied alone.

The above discussion is based on the noise-free assumption. Because the noise is spatially and temporally white, the averaging process in the presence of noise could also provide an enhancement of the signal-to-noise ratio (SNR) in $\bar{W}(t,f)$ by a factor of M over the WVD of the single sensor. This amounts to increasing the synthesis robustness with respect to noise, which becomes important in the environment where the desired signals are submerged within the noise. Therefore the benefits of the proposed method is two-fold, reduction of the signal crossterms as well as the additive noise level.

III. Signal Synthesis

A. Signal Synthesis Using Linear TFDs

Unlike bilinear approaches, methods that use linear filtering to recover or estimate the time-varying signal could avoid the troublesome crossterms. There are several methods in linear synthesis approach. The most common and straightforward one is by short-time Fourier Transform (STFT) [9]. If the masked STFT is still a valid transform, then the signal could be recovered perfectly. If not, usually the MSE criterion is applied to obtain

the most optimal solution. Another linear approach in signal synthesis is based on wavelet transformation (WT) [11]. Related works could also be found out in [9], [11].

Although capable of solving many nonstationary problems, the linear approach is not always an appropriate technique to restore the signals with fast-changing t-f signatures, e.g., the linear polynomial phase signals. A natural and reliable approach for FM signals is through the bilinear t-f transforms, typically, the WVD.

B. Analogy of the Array WVD Signal Synthesis

The proposed array averaging technique structurely resembles the "weighted model" introduced in [7], and is given by

$$\hat{\mathrm{W}}(t,f) = \sum_{i=1}^{L}\sum_{j=1}^{L} \alpha_{ij} \mathrm{W}_{\mathrm{s}_i\mathrm{s}_j}(t,f) \tag{25}$$

where α_{ij} is the weighting factor for the signal sources' auto- or cross-terms $\mathrm{W}_{\mathrm{s}_k\mathrm{s}_l}(t,f)$. $\hat{\mathrm{W}}(t,f)$ denotes the weighted t-f distribution and is similar in structure to the array averaged WVD $\bar{\mathrm{W}}(t,f)$ defined in (12). The key difference between equation (25) and (12) is that, in the underlying problem, we do not intentionally select the weight factors β_{ij} in (12). They evolve naturally from the inner products of the associated SSs embedded in the matrix of $\mathbf{\Gamma}$ and are generated without any "human intervention" through the process of array averaging.

To synthesize the signal from the averaged WVD, we apply transformation to the mixing matrix as

$$\mathbf{B} = \mathbf{A} \cdot diag(\mathrm{d}_i, \ldots, \mathrm{d}_L) \tag{26}$$

where $\mathrm{d}_i = \|\mathrm{s}_i(t)\|, \quad i = 1, \ldots, L$. This is equivalent to normalizing the signal power and exchanging the scalar factor in $\mathrm{s}_i(t)$ to the mixing matrix $\mathbf{B}$. Next, we perform the following singular value decomposition (SVD),

$$\mathbf{U}^H(\mathbf{B})\mathbf{V} = diag(\lambda_1, \ldots, \lambda_L) \tag{27}$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_L \geq 0$, and $\mathbf{U}$ and $\mathbf{V}$ are $M \times M$ and $L \times L$ unitary matrices, respectively. Denote $\mathbf{\Gamma}' = \frac{1}{M}\mathbf{B}^H\mathbf{B}$, it is straightforward to show that

$$\mathbf{V}^H(\mathbf{\Gamma}')\mathbf{V} = diag(\frac{\lambda_1^2}{M}, \ldots, \frac{\lambda_L^2}{M}) \tag{28}$$

Following the same procedures in [7],we could derive the result of the synthesized signal from the averaged WVD as,

$$\mathrm{s}_{opt}(t) = \sum_{k=1}^{L} \gamma_k \mathrm{s}_k(t) \tag{29}$$

with

$$\gamma_k = \frac{e^{j\phi}\lambda_k \mathrm{v}_{1k}}{\sqrt{M}\mathrm{d}_k}, \tag{30}$$

where v_{1k} is the kth element of matrix $\mathbf{V}$'s first column $\mathbf{v}_1$, that is, the eigenvector of $\mathbf{\Gamma}'$ corresponding to the largest eigenvalue λ_1^2/M. From (29), and drawing an analogy to the weighted model in [7], it is clear that the synthesized signal using the array averaging technique is a linear combination of the source signals $\mathrm{s}_k(t)$. The combination coefficients γ_k are irrelevant to the signals' t-f characteristics and depend only on the formation of matrix $\mathbf{\Gamma}'$ which is only a variation of $\mathbf{\Gamma}$.

Equation (29) shows theoretical results of signal synthesis based on the averaged WVD. Without any mask that distinguishes the different signal auto-terms, the synthesized signal is generally a combination of the source signals $\mathrm{s}_k(t)$, which is no different from the signal received at a single antenna sensor. In the special case where there is no cross-terms present, (29) gives a solution of a single source signal with the maximum energy [7].

C. Single Signal Synthesis from Averaged WVD

In order to recover the sources respective signals without crossterms contamination, which become mostly suppressed by the array averaging process, an appropriate mask is often placed on the autoterms of each source in the t-f domain.

Practically, the WVD-based synthesis techniques could be found in [6], [7], [8], [9]. In this paper, we apply the method of *extended discrete-time Wigner distribution* (EDTWD), introduced in [9], to the output of array averaged WVD. The advantage of using the EDTWD lies in the fact that it does not require *a priori* knowledge of the source waveform, and thereby avoids the problem of matching the two "uncoupled" vectors (even-indexed and odd-indexed vectors).

The overall synthesis procedure is summarized in the following steps.

1. Given the received data of the ith sensor $\mathrm{x}_i(t)$, compute the EDTWD

$$\mathrm{W}_{\mathrm{x}_i\mathrm{x}_i}(t,f) = \sum_{k:t+\frac{k}{2}\in Z} \mathrm{x}_i(t+\frac{k}{2})\mathrm{x}_i{}^*(t-\frac{k}{2})e^{-j2\pi kf}, \quad t = 0, \pm 0.5, \pm 1, \ldots. \tag{31}$$

2. Apply the averaging process, that is, summing the EDTWD across the array

$$\bar{\mathrm{W}}(t,f) = \frac{1}{M}\sum_{k=1}^{M} \mathrm{W}_{\mathrm{x}_k\mathrm{x}_k}(t,f). \tag{32}$$

3. Place an appropriate t-f mask on $\bar{\mathrm{W}}(t,f)$ such that only the desired signal autoterms are retained.

4. Take the IFFT of the masked WVD $\bar{\mathrm{W}}(t,f)$

$$p(t,\tau) = \int \bar{\mathrm{W}}(t,f)e^{j2\pi\tau f}df. \tag{33}$$

5. Construct the matrix $\mathbf{Q} = [\mathrm{q}_{ij}]$ with

$$\mathrm{q}_{ij} = p\left(\frac{i+j}{2}, i-j\right). \tag{34}$$

6. Apply eigendecomposition to the matrix $[\mathbf{Q}+\mathbf{Q}^H]$ and obtain the maximum eigenvalue λ_{max} and the associated eigenvector $\mathbf{u}$. The desired signal is given by

$$\hat{\mathrm{s}}_{opt} = e^{j\phi}\sqrt{2\lambda_{max}}\ \mathbf{u}. \tag{35}$$

7. Repeat step 3 through 6 until all source signals $\hat{\mathrm{s}}_1(t), \hat{\mathrm{s}}_2(t), \ldots, \hat{\mathrm{s}}_L(t)$ are retrieved.

D. Array Matrix Estimation

Upon synthesizing all the source signals, we could utilize these signal waveforms to estimate the mixing, or array, matrix $\mathbf{A}$ through the MSE criterion,

$$\varepsilon = \sum_{t=1}^{N} \left\| \mathbf{x}(t) - \mathbf{A}\hat{\mathbf{s}}(t) \right\|^2 = \sum_{t=1}^{N} \left(\mathbf{x}(t) - \sum_{i=1}^{L} \mathbf{a}_i \hat{\mathrm{s}}_i(t) \right)^H \left(\mathbf{x}(t) - \sum_{j=1}^{L} \mathbf{a}_j \hat{\mathrm{s}}_j(t) \right). \tag{36}$$

The solution is to find the matrix that minimizes the cost function ε, that is,

$$\hat{\mathbf{A}} = \arg\min_{\mathbf{A}} \varepsilon. \tag{37}$$

The necessary conditions for (37) is

$$\begin{aligned} \frac{\partial \varepsilon}{\partial \mathbf{a}_k} &= -\sum_{t=1}^{N} \left[\hat{\mathrm{s}}_k^*(t) \left(\mathbf{x}(t) - \sum_{i=1}^{L} \mathbf{a}_i \hat{\mathrm{s}}_i(t) \right) \right] \\ &= -\hat{\mathbf{r}}_k + \sum_{i=1}^{L} \mathbf{a}_i \hat{\mathbf{R}}_{ik} \\ &= \mathbf{0}, \quad k = 1, \ldots, L. \end{aligned} \tag{38}$$

where

$$\hat{\mathbf{R}} = \sum_{t=1}^{N} \hat{\mathbf{s}}(t) \hat{\mathbf{s}}^H(t) \tag{39}$$

represents the estimated signal source covariance matrix, and $\hat{\mathbf{r}} = [\hat{\mathbf{r}}_1, \ldots, \hat{\mathbf{r}}_L]$, with

$$\hat{\mathbf{r}}_i = \sum_{t=1}^{N} \mathbf{x}(t) \hat{\mathrm{s}}_i^*(t) \tag{40}$$

is the correlation vector between the data vector received across the array and the ith source signal $\hat{\mathrm{s}}_i(t)$. In equations (38-40), "$\wedge$" signifies the fact that we deal with estimated variables. The matrix form for all the L equations expressed in (38) is

$$-\hat{\mathbf{r}} + \mathbf{A}\hat{\mathbf{R}} = \mathbf{0}. \tag{41}$$

The estimate of the array matrix is

$$\hat{\mathbf{A}} = \hat{\mathbf{r}}\hat{\mathbf{R}}^{-1}. \tag{42}$$

It is evident from the implementation of the proposed algorithm, that the spatial information needs to be compromised in the first phase to permit crossterms and noise suppressions. The mixing matrix $\mathbf{A}$ could be estimated only after the complete retrieval of the signal source waveforms. This is fundamentally different from other algorithms that combine array signal processing with conventional bilinear distributions, e.g., the spatial time-frequency distribution (STFD) in which the mixing matrix estimation precedes the estimation of the source signals and is provided using relationship

$$\mathbf{D_{xx}}(t,f) = \mathbf{A}\mathbf{D_{ss}}(t,f)\mathbf{A}, \tag{43}$$

where $\mathbf{D_{ss}}(t,f)$ is the signal TFD matrix whose entries are the auto- and cross-TFD's of the sources and $\mathbf{D_{xx}}(t,f)$ is the data STFD matrix. In the STFD-based source separation, the estimate of $\mathbf{A}$ is provided using whitening, followed by joint diagonalization of $\mathbf{D_{xx}}(t,f)$ for $(t,f \in$ autoterm regions). This estimate is then included to obtain the source signals using Pseudo-inverse of $\mathbf{A}$.

A hybrid technique based on both array averaging and STFD can be adopted The array averaging of WVDs is first performed to offer a good estimate of t-f signatures of source signals through cross-term suppression properties. Once the auto-source WVDs are determined, we could then construct the STFD matrices and recover the synthesized signal waveforms, as well as the mixing matrix based on equation (42).

E. Computational Cost

To compare the computational cost of the proposed method and STFD, we use the number of complex multiplications as the evaluation criterion. For the array averaged WVD, the computational cost is shown to be (Appendix A)

$$N_{aa} \approx (1 + L + 4\log_2 2N) \cdot O(N^2), \tag{44}$$

where the operand $O(\cdot)$ denotes the order of "$\cdot$". Equation (44) shows that the computation cost required for the proposed method is almost proportional to $O(N^2)$.

The computational cost for a typical STFD process is given as (also see Appendix A)

$$N_{STFD} \approx \log_2 N \cdot O(N^2). \tag{45}$$

Thus the array averaging technique requires more computations than the STFD.

F. Signal Synthesis with Overlapping T-F Signatures

The procedures we have discussed is appropriate to synthesize the signal waveform whose t-f signatures are distinct. In this case, the masked t-f region always contains the autoterm of the desired source signal with the influence from other sources often negligible. However, if the source t-f signatures overlap, the mask is deemed to capture undesired autoterms. This problem cannot be mitigated by spatial averaging of TFDs and a modification of the proposed method is in order.

Assume that upon implementing the synthesis process described in section III.D, we obtain the estimate of the mixing matrix $\hat{\mathbf{A}}$. Since there are interfering signal autoterms from other sources, $\hat{\mathbf{A}}$ should be considered different from $\mathbf{A}$. We use $\hat{\mathbf{A}}$ to construct a beamformer applied to the data received across the array (assume the noise-free scenario). That is,

$$\mathbf{z}(t) = \frac{1}{M}\hat{\mathbf{A}}^H\mathbf{x}(t) = \frac{1}{M}\hat{\mathbf{A}}^H\mathbf{A}\mathbf{s}(t). \tag{46}$$

where $\mathbf{z}(t) = [\mathrm{z}_1(t), \ldots, \mathrm{z}_L(t)]$ is a $L \times 1$ vector. Clearly,

$$\mathrm{z}_k(t) = \left(\frac{1}{M}\hat{\mathbf{a}}_k^H\mathbf{a}_k\right)\mathrm{s}_k(t) + \sum_{l\neq k}^{L}\left(\frac{1}{M}\hat{\mathbf{a}}_k^H\mathbf{a}_l\right)\mathrm{s}_l(t). \tag{47}$$

It is expected that $\hat{\mathbf{a}}_k$ is a perturbed version of $\hat{\mathbf{a}}_k$. With the approximations

$$\left|\frac{1}{M}\hat{\mathbf{a}}_k^H\mathbf{a}_k\right| \approx \beta_{kk} = 1 \tag{48}$$

and

$$\left|\frac{1}{M}\hat{\mathbf{a}}_k^H\mathbf{a}_l\right| \approx \beta_{lk} \ll 1, l \neq k. \tag{49}$$

then, the WVD of $z_k(t)$ is given by

$$W_{z_k z_k}(t,f) \approx W_{s_k s_k}(t,f) + \sum_{i=1}^{L} \sum_{(j=1, j\neq i)}^{L} \beta_{ik}\beta_{jk}^{*} W_{s_i s_j}(t,f). \tag{50}$$

for $j \neq i$, $\beta_{ik}\beta_{jk}^{*} \ll 1$. This shows that in equation (50), except the kth auto-source term, all other terms, either auto- or cross-source terms, are significantly reduced from $W_{z_k z_k}(t,f)$. In the case of ULA, the suppression of those terms are at least 13dB for large value of M. The suppression of the autoterms other than source k is $|\beta_{ik}|^2$, which is more than 26dB down from the kth source. Therefore, the effect of the overlapping autoterms from other sources becomes negligible. If we apply the steps (3-8) of the synthesis procedure of section III-D using the improved WVD in (50), the synthesized signal will be enhanced.

IV. Simulation Results

In this section, we provide computer simulations to demonstrate the improvement gained by the proposed technique in the reduction or elimination of crossterms. Specifically, we examine the effect of array averaging on the retrieval and separation of the nonstationary signals impinging on the multi-sensor array. In all the simulations presented below, we consider several signals incident on an eight-sensor ULA ($M = 8$) with inter-element spacing of half-wavelength. The additive noise is zero mean, Gaussian distributed, spatially and temporally white. The length of the signal sequence is set to $N = 128$.

Moreover, we use the same performance index applied in [12] and [14] to evaluate the performance of the proposed technique

$$I_{pq} = E\left| \left(\hat{\mathbf{A}}^{\#}\mathbf{A}\right)_{pq} \right|^2, \tag{51}$$

where the superscript $^{\#}$ denotes the pseudo-inverse. Equation (51) defines the interference-to-signal ratio (ISR). Thus, I_{pq} measures the ratio of the power of the interference of qth source signal to the power of the pth source signal. For large enough N, we have $I_{pq} \approx 0$

for $p \neq q$. We also apply the global rejection level to evaluate the overall performance of the proposed method

$$I_{perf} = \sum_{q \neq p} I_{pq}. \tag{52}$$

In the first example, three chirp signals, $s_1(t)$, $s_2(t)$ and $s_3(t)$, arrive at the array with AOAs of $-20°$, $0°$ and $20°$, with the respective start and end frequencies given by $(0.9\pi, 0.5\pi)$, $(0.66\pi, 0.26\pi)$, and $(0.5\pi, 0.1\pi)$. In the t-f plane, the source signals have parallel signatures, emulating a multipath environment. The crossterm of $s_1(t)$ and $s_3(t)$ also forms a chirp-like crossterm structure whose frequency starts from 0.7π and ends with 0.3π, and therefore lies closely to the t-f signature of $s_2(t)$. Fig. 1 depicts the WVD of the signals at the reference sensor (sensor #1) for the case of noise-free environment. It is clear that the t-f signature of all signal autoterms and crossterms are parallel in the t-f domain. The crossterms produced from the three source signals are even more dominant than the source autoterms. In the single sensor receiver, it becomes difficult to distinguish the source autoterms from the crossterms without any *a priori* knowledge of the sources. From the above AOAs, we obtain

$$\left|\frac{1}{8}\mathbf{A}^H\mathbf{A}\right| = \begin{bmatrix} 1 & 0.2236 & 0.1048 \\ 0.2236 & 1 & 0.2236 \\ 0.1048 & 0.2236 & 1 \end{bmatrix}. \tag{53}$$

The off-diagonal elements are small ($< -13dB$), compared to the matrix diagonal entries, indicating that the sources spatial signatures are weakly correlated, and the array averaging process could result in a substantial reduction in the crossterms. Fig. 2 shows the corresponding array-averaged WVD. Due to the reduction in cross-terms by more than 13dB, the t-f signatures of the sources are distinctively exhibited in the plots. Explicitly, the crossterm from $s_1(t)$ and $s_3(t)$ ceased to become an interfering factor in identifying the adjacent signal source $s_2(t)$. In effect, averaging the WVDs across the array has significantly reduced the crossterms, whereas the three signals' autoterms have remained

intact.

Next, we add 5dB noise to the data at each array sensor so that the input SNR is -5dB. Fig. 3 and 4 depict both the reference-sensor WVD and the array-averaged WVD. It is evident that the noise obscures both the signal autoterms and crossterms of the WVD at a single sensor. It is difficult, therefore, to retrieve the desired signal if we only synthesize from a single sensor.

Upon averaging, both noise and crossterms are sufficiently reduced to clearly manifest the individual source t-f signature, and the signals could be individually recovered if we place the appropriate masks in the t-f region. Fig. 5 and Fig. 6 shows the WVD of the synthesized signal $s_2(t)$ using the array averaging and STFD techniques, respectively. Fig. 7 displays the real parts of the original signal $s_2(t)$, the STFD-recovered $\hat{s}_2(t)$, and $\hat{s}_2(t)$ synthesized by the proposed method. It is clear that the result from the array averaging technique is closer to the original signal than the recovered signal from the STFD-based method. We also plot the global rejection level I_{perf} versus the input SNR in Fig. 8. The input SNR takes values from -10dB to 20dB. Both the STFD-based and the array averaging-based techniques are used to compute the empirical I_{perf} defined in equation (52). Increasing the SNR certainly improves the performance for both methods, and simulations show that the STFD-based method is outperformed by the array averaging technique, which is consistent with the results given in Fig. 7.

In the second example, we use two chirp signals with overlapping autoterms. The signals are from AOAs of $-20°$ and $20°$ with start and end frequencies of $(0.7\pi, 0.3\pi)$ and $(0.3\pi, 0.7\pi)$, respectively. There is no additive noise in this example. Fig. 9 shows the WVD of data from the reference sensor #1. The two signal autoterms overlap, and their cross source terms could also be clearly noticed. The array averaged WVD is plotted in Fig. 10. Using the conclusions derived in section II, we expect that the cross-source terms would be suppressed by about 19dB after the array averaging process. This is supported by the plots in Fig. 10. To synthesize the signal, we place the mask along each

t-f signature. Any reasonable selection of the mask inevitably includes components from the other source. Therefore, each signal synthesized following the procedures described in section III-D is, in essence, corrupted by the other signal. Fig. 11(b) depicts the WVD of one synthesized but corrupted waveform, compared to the WVD from the original source, which is shown in Fig. 11(a). By further implementing the beamformer and synthesis procedures from section III-E, we could obtain less noisy waveform. The WVD of the improved synthesized signal is shown in Fig. 11(c). The power leakage in Fig. 11(b) almost disappears in Fig. 11(c). The corresponding global rejection levels are calculated and equal to 0.02652 and 0.00159 for the respective cases, showing the important role of the proposed beamformer approach.

V. Conclusion

A two-step synthesis technique using bilinear distributions was proposed for multi-sensor receivers. The first step is to average the Wigner-Ville distributions of the sensor data across the array. This averaging process allows the distinction in the spatial structures of the sources to play a key role in improving their time-frequency representations. This improvement is manifested in the reduction of the noise floor and mitigation of cross-terms in the t-f domain. The second step is to apply well-known bilinear synthesis methods to the averaged WVD. It was shown that the proposed synthesis approach is fundamentally different form the one recently devised using spatial time-frequency distributions. In the latter, the source spatial signatures need to be first estimated before the sources could be separated. The main attraction of the proposed approach is that it naturally extends bilinear signal synthesis to array processing. In doing so, it capitalizes on the spatial dimension to reduce the cross-terms without smearing the auto-terms, which can not be achieved using the t-f smoothing operation via reduced interference distributions.

Appendix A: Computation for STFD-Based and Array Averaging-Based Techniques

We first derive the computational requirement for a single EDTWD counting the complex multiplications. It is clear from equation (42) that for a sequence $\mathrm{y}(t)$ $(t = 0, 1, \ldots, N-1)$, the EDTWD would generate a $(2N-1) \times (2N-1)$ two-dimension time-frequency distribution $\mathbf{W}_{yy}$. The computation cost is determined by constructing a $(2N-1) \times (2N-1)$ matrix whose element are given by

$$\mathrm{W}(t,k) = \mathrm{x}_i(t+\frac{k}{2})\mathrm{x}_i{}^*(t-\frac{k}{2})e^{-j2\pi kf}, \quad t = 0, 0.5, 1, \ldots, N-1, k = 0, \pm 1, \ldots, \pm(N-1). \tag{A.1}$$

The number of complex multiplication involved in (A.1) is

$$n_1 \;=\; 1+2+\ldots+(N-1)+N+(N-1)+\ldots+1 = N^2. \tag{A.2}$$

The EDTWD could be obtained simply by calculating the fast Fourier transform (FFT) with respect to the k parameter in $\mathrm{W}_1(t,k)$. Assume N_{FFT} is the computation cost for a sequence of length $(2N-1)$, we have $N_{FFT} \approx 2N \times \log_2 2N$. Therefore the computation cost required in a single EDTWD is

$$n_2 \;=\; (2N-1) \times N_{FFT} \approx 4N^2 \log_2 2N. \tag{A.3}$$

Other costs are N_{FFT} for IFFT, $L \cdot O(N^2)$ (according to [16]) for eigendecomposition to recover source waveforms, $L(L+1)/2N + LMN + O(L^2)$ for recovery of mixing matrix. The total cost is computed by summing the above results. Since the number of samples N is often much greater than other parameters, we neglect all items but $O(N^2)$, yielding

$$N_{aa} = n_1 + n_2 + N_{FFT} + L \cdot O(N^2) \approx (1 + L + 4\log_2 2N) \cdot O(N^2). \tag{A.4}$$

The computation requirements of STFD could be calculated in two parts. The first part is related to extractions of the t-f signature. This is achieved by computing the

WVD of the data from a single reference sensor and selecting the t-f points that are associated with the different signal sources. Similar to EDTWD, the required cost for this process is $\log_2 N \cdot O(N^2)$. The second part involves the constructing of the STFD matrices and recovery of the signal waveforms. It is straightforward to show the respective computation cost is: Estimation of the auto correlation matrix: $NM^2/2$; Computation of the whitening matrix: $O(M^3)$; Whitening of the data: NLM; Estimation of the STFDs: $\alpha_T L(L-2)/2$; Joint Diagonalization: $O(KL^3)$; Separation: LMN. The variables L, M, and N are the number of sources, sensors and samples, respectively, whereas K is the number of the chosen t-f points. α_T is the cost of one classical TFD. In a typical scenario, $L, M, K, \alpha_T \ll N$, Therefore the computation cost for STFD is

$$N_{STFD} \approx \log_2 N \cdot O(N^2). \tag{A.5}$$

References

[1] L. Cohen, "Time-frequency distributions – a review," *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, July 1989.

[2] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing*, S. Haykin ed., vol. 1, Chapter 9, pp. 21–68, Englewood Cliffs, NJ: Prentice-Hall, 1990.

[3] F. Hlawatsch and G. Boudreaux-Bartles, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, no. 2, pp. 21–68, April 1992.

[4] L. Cohen, *Time–Frequency Analysis*, Englewood Cliffs, NJ: Prentice Hall, pp. 162–167, 1995.

[5] S. Qian and D. Chen, *Joint Time-Frequency Analysis*, Englewood Cliffs, NJ: Prentice-Hall, 1996.

[6] G. Boudreaux–Bartels and T. Parks, "Time–varying filtering and signal estimation using Wigner distribution synthesis techniques," *IEEE Trans. ASSP*, vol. ASSP–34, no. 3, pp. 442–451, June 1986.

[7] W. Krattenthaler and F. Hlawatsch, "Time–frequency design and processing of signals via smoothed Wigner distributions," *IEEE Trans. Signal Processing*, vol. 41, no. 1, pp. 278–287, Jan. 1993.

[8] F. Hlawatsch and W. Krattenthaler, "Bilinear signal synthesis," *IEEE Trans. Signal Processing*, vol. 40, no. 2, pp. 352–363, Feb. 1992.

[9] J. Jeong and W. Williams "Time–varying filtering and signal synthesis," in *Time–Frequency Signal Analysis — Methods and Applications*,
B. Boashash ed., Longman Cheshire Pty, 1995.

[10] J. Jeong and W. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Processing*, vol. 40, pp. 402–412, Feb. 1992.

[11] Daubechies, "The wavelet transform, time–frequency localization and signal analysis," *IEEE Trans. Info. Theory*, vol. IT–36, no. 5, pp. 961–1005, Sep. 1993.

[12] A. Belouchrani and M. Amin, "Blind source separation based on time–frequency signal representation," *IEEE Trans. Signal Processing*, vol. 46, pp. 2888–2897, Nov. 1998.

[13] Y. Zhang, W. Mu, and M. Amin, "Subspace analysis of spatial time-frequency distribution matrices," *IEEE Trans. Signal Processing*, vol. 49, no. 4, April 2001.

[14] Y. Zhang and M. Amin, "Blind separation of sources based on their time-frequency signatures," in *Proc. IEEE ICASSP*, Istanbul, Turkey, pp. 3132–3135, June 2000.

[15] C. Balanis, *Antenna Theory – Analysis and Design*, John Wiley & Sons, pp. 257–261, 1997.

[16] G. Xu and T. Kailath, "Fast subspace decomposition," *IEEE Trans. Signal Processing*, vol. 42, pp. 539-551, Mar. 1994.

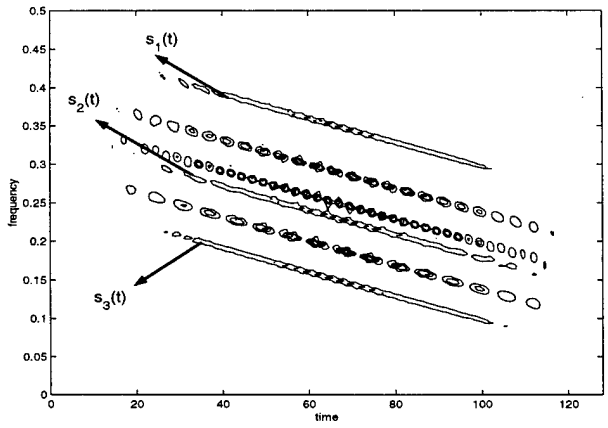


Fig. 1. WVD in noise-free case at reference sensor.

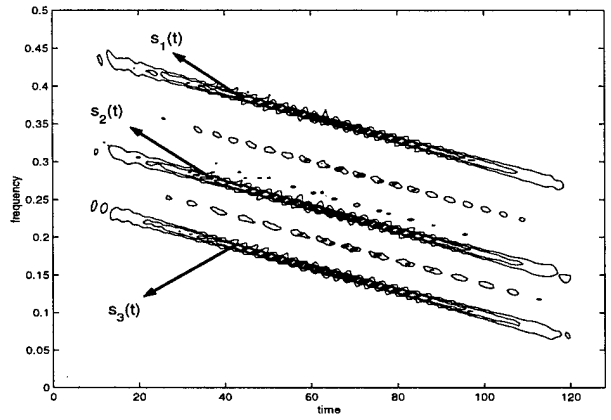


Fig. 2. Array-averaged WVD in noise-free environment.

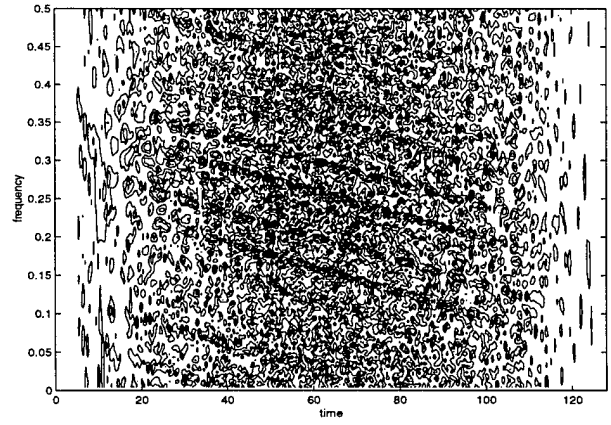


Fig. 3. WVD of the corrupted signals at reference sensor.

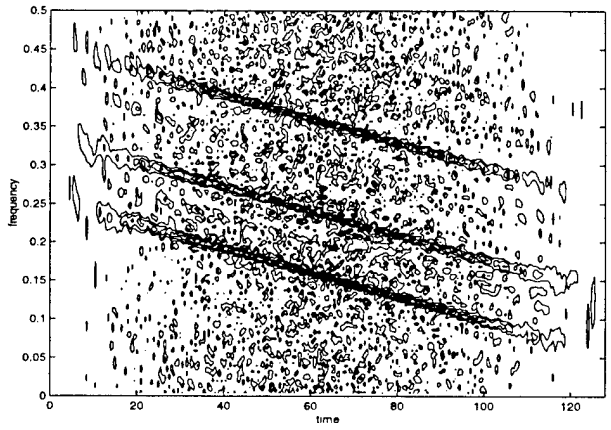


Fig. 4. Array-averaged WVD of the corrupted signals.

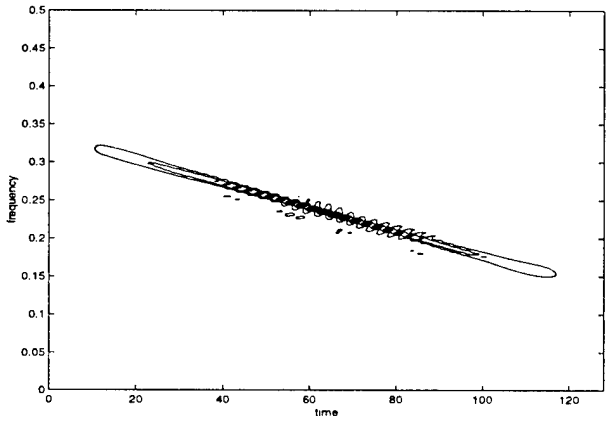


Fig. 5. WVD of synthesized $\hat{s}_2(t)$ using array-averaging.

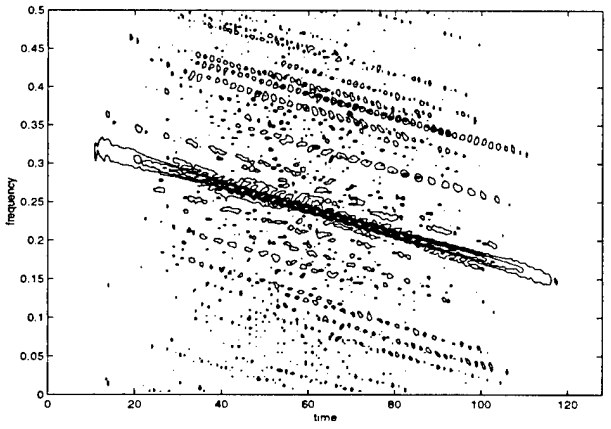


Fig. 6. WVD of synthesized $\hat{s}_2(t)$ using STFD.

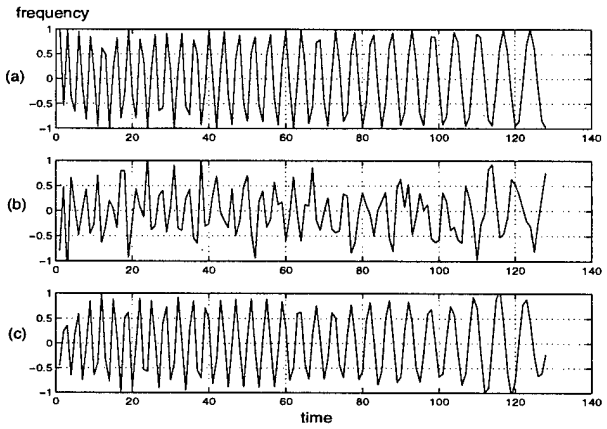


Fig. 7. (a) Real part of original $s_2(t)$ (top); (b) Real part from the STFD-recovered $\hat{s}_2(t)$ (middle); (c) Real part from the array averaged $\hat{s}_2(t)$ (bottom).

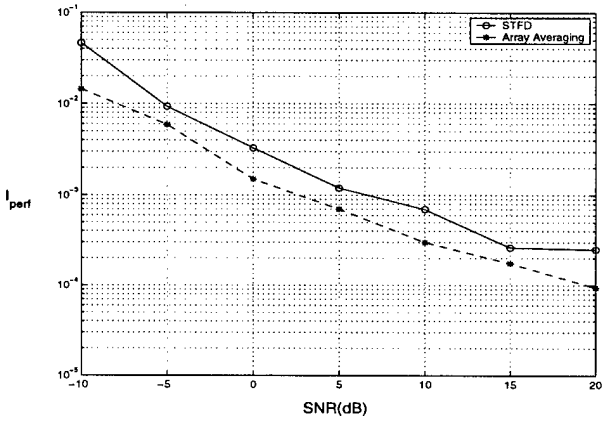


Fig. 8. Global rejection level versus input SNR. (o: by STFD; *: by array averaging).

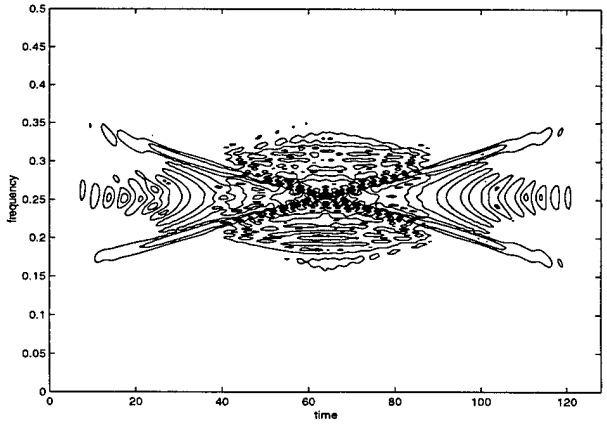


Fig. 9. WVD of the two overlapping signals from a reference sensor.

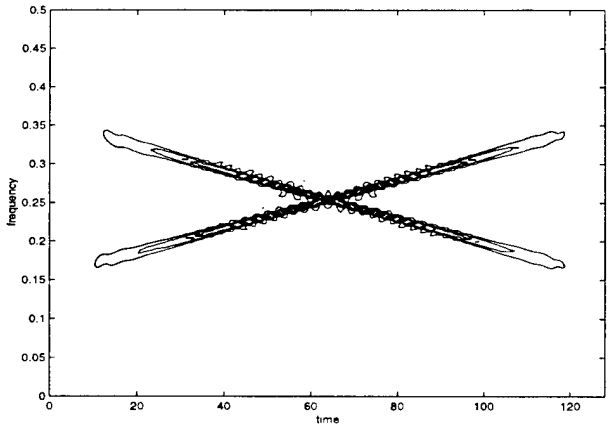


Fig. 10. Array averaged WVD.

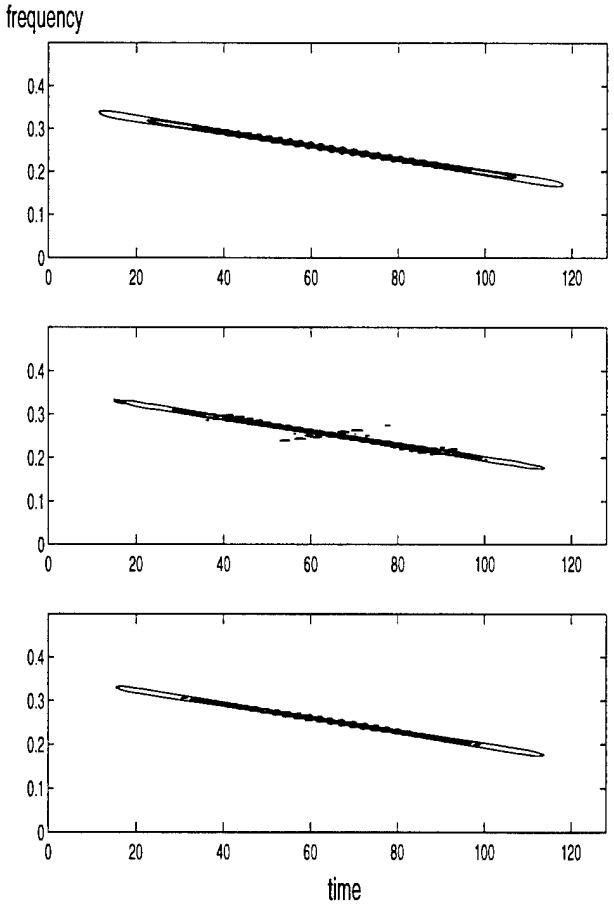


Fig. 11. WVD of: (a) original signal (top); (b) synthesized signal from array averaging (middle); (c) synthesized signal based on beamforming (bottom).

Time-frequency maximum likelihood methods for direction finding

Yimin Zhang, Weifeng Mu, and M. G. Amin
Department of Electrical and Computer Engineering,
Villanova University, Villanova PA 19085 USA

Abstract

This paper proposes a novel time-frequency maximum likelihood (t-f ML) method for direction-of-arrival (DOA) estimation for non-stationary signals impinging on a multi-sensor array receiver, and compares this method with conventional maximum likelihood DOA estimation techniques. Time-frequency distributions localize the signal power in the time-frequency domain, and as such enhance the effective SNR, leading to improved DOA estimation. The localization of signals with different time-frequency signatures permits the division of the time-frequency domain into smaller regions, each containing fewer signals than those incident on the array. The reduction of the number of signals within different time-frequency regions not only reduces the required number of sensors, but also decreases the computational load in multi-dimensional optimizations. Compared to the recently proposed time-frequency MUSIC (t-f MUSIC), the proposed t-f ML method can be applied in coherent environments, without the need to perform any type of preprocessing that is subject to both array geometry and array aperture.

Keywords

Time-frequency distribution, direction finding, maximum likelihood, spatial time-frequency distribution, array processing.

This work was supported by the Office of Naval Research under Grant N00014-98-1-0176.

I. Introduction

The localization of spatial sources by passive sensor array is one of the important problems in radar, sonar, radio-astronomy, and seismology. So far, numerous methods have been proposed for direction finding, most of which are based on the estimates of the data covariance matrix. Among these methods, the maximum likelihood (ML) technique was one of the first to be devised and investigated [1]. It has a superior performance compared to other methods, particularly when the input signal-to-noise ratio (SNR) is low, the number of data samples is small, or the sources are highly correlated [2]. Therefore, despite its complexity, the ML technique remains of practical interest.

The evaluation of quadratic time-frequency distributions of the data snapshots across the array yields what is known as spatial time-frequency distributions (STFDs) [3], [4]. STFD techniques are most appropriate to handle sources of nonstationary waveforms. STFDs spread the noise power while localizing the energy of the impinging signals in the time-frequency domain. This property leads to increasing the robustness of eigenstructure signal and noise subspace estimates with respect to the channel and receiver noise, and hence improves spatial resolution performance.

In this paper, we propose the time-frequency maximum likelihood (t-f ML) method for direction finding and provide the analysis that explains its performance. It is shown that the superior performance of the t-f ML method relative to other methods is attributed to the following three fundamental reasons: 1) Time-frequency distributions localize the signal power in the time-frequency domain, and as such enhance the effective SNR and improve the direction-of-arrival (DOA) estimation. 2) The localization of signals with different time-frequency signatures permits the division of the time-frequency domain into smaller regions, each containing fewer signals than those incident on the array. The reduction of the number of signals within different time-frequency regions relaxes the condition on the size of the array aperture as well as simplifies the multidimensional optimization estimation procedure. 3) Compared with the previously proposed time-frequency MUSIC (t-f MUSIC), the t-f ML method can be applied when the signal arrivals are highly correlated, whereas the t-f MUSIC algorithm cannot do so without some sort of preprocessing, such as spatial smoothing.

This paper is organized as follows. In Section 2, the signal model is established, and a brief review of the spatial time-frequency distributions is given. In Section 3, we discuss the SNR enhancement based on time-frequency distributions and its effect on the signal and noise subspace estimates using STFD matrices. The subspace estimates obtained from the STFD matrices are more robust to SNR and angular separation compared to those obtained from data covariance matrices. Section 4 presents the t-f ML and shows its performance in time-varying environments.

II. Background

A. Signal model

In narrowband array processing, when n signals arrive at an m-element array, the linear data model

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{A}(\mathbf{\Theta})\mathbf{d}(t) + \mathbf{n}(t) \tag{1}$$

is commonly assumed, where the $m \times n$ spatial matrix $\mathbf{A}(\mathbf{\Theta}) = [\mathbf{a}(\theta_1), \cdots, \mathbf{a}(\theta_n)]$ represents the mixing matrix or the steering matrix, and $\mathbf{a}(\theta_i)$ are the steering vectors corresponding to angle of arrival θ_i. Due to the mixture of the signals at each sensor, the elements of the $m \times 1$ data vector $\mathbf{x}(t)$ are multicomponent signals, whereas each source signal $d_i(t)$ of the $n \times 1$ signal vector $\mathbf{d}(t)$ is often a monocomponent signal. $\mathbf{n}(t)$ is an additive noise vector whose elements are modeled as stationary, spatially and temporally white, zero-mean complex random processes, independent of the source signals. That is,

$$E[\mathbf{n}(t+\tau)\mathbf{n}^H(t)] = \sigma\delta(\tau)\mathbf{I} \text{ and } E[\mathbf{n}(t+\tau)\mathbf{n}^T(t)] = \mathbf{0} \text{ for any } \tau \tag{2}$$

where $\delta(\tau)$ is the Kronecker delta function, $\mathbf{I}$ denotes the identity matrix, σ is the noise power at each sensor, superscript ^{H} and ^{T} respectively denote conjugate transpose and transpose, and $E(\cdot)$ is the statistical expectation operator.

In equation (1), it is assumed that the number of sensors is greater than the number of sources, i.e., $m > n$, and the number of snapshots is greater than the number of array sensors, i.e., $N > m$. We also assume that matrix $\mathbf{A}$ is full column rank, which implies that the steering vectors corresponding to n different angles of arrival are linearly independent.

Under the above assumptions, the correlation matrix is given by

$$\mathbf{R_{xx}} = E[\mathbf{x}(t)\mathbf{x}^H(t)] = \mathbf{A}(\Theta)\mathbf{R_{dd}}\mathbf{A}^H(\Theta) + \sigma\mathbf{I}, \tag{3}$$

where $\mathbf{R_{dd}} = E[\mathbf{d}(t)\mathbf{d}^H(t)]$ is the signal correlation matrix. For notational convenience, we drop the argument Θ and simply use $\mathbf{A}$ instead of $\mathbf{A}(\Theta)$. If $\hat{\Theta}$ is an estimate of Θ, then we also use $\hat{\mathbf{A}}$ instead of $\mathbf{A}(\hat{\Theta})$.

Let $\lambda_1 > \lambda_2 > \cdots > \lambda_n > \lambda_{n+1} = \lambda_{n+2} = \cdots = \lambda_m = \sigma$ denote the eigenvalues of $\mathbf{R_{xx}}$. The unit-norm eigenvectors associated with $\lambda_1, ..., \lambda_n$ constitute the columns of matrix $\mathbf{S} = [\mathbf{s}_1, ..., \mathbf{s}_n]$, and those corresponding to $\lambda_{n+1}, ..., \lambda_m$ make up matrix $\mathbf{G} = [\mathbf{g}_1, ..., \mathbf{g}_{m-n}]$. Since the columns of $\mathbf{A}$ and $\mathbf{S}$ span the same subspace, then $\mathbf{A}^H\mathbf{G} = \mathbf{0}$.

In practice, $\mathbf{R_{xx}}$ is unknown, and therefore should be estimated from the available data samples (snapshots) $\mathbf{x}(i)$, $i = 1, 2, ..., N$. The estimated correlation matrix is given by

$$\hat{\mathbf{R}}_{\mathbf{xx}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}(i)\mathbf{x}^H(i). \tag{4}$$

Let $\{\hat{\mathbf{s}}_1, ..., \hat{\mathbf{s}}_n, \hat{\mathbf{g}}_1, ..., \hat{\mathbf{g}}_{m-n}\}$ denote the unit-norm eigenvectors of $\hat{\mathbf{R}}_{\mathbf{xx}}$, arranged in the descending order of the associated eigenvalues, and let $\hat{\mathbf{S}}$ and $\hat{\mathbf{G}}$ denote the matrices made of the set of vectors $\{\hat{\mathbf{s}}_i\}$ and $\{\hat{\mathbf{g}}_i\}$, respectively. The statistical properties of the eigenvectors of the sample covariance matrix $\hat{\mathbf{R}}_{\mathbf{xx}}$ for signals modeled as independent processes with additive white noise is given in [6].

In this paper, we focus on frequency-modulated (FM) signals, modeled as

$$\mathbf{d}(t) = [d_1(t), ..., d_n(t)]^T = [D_1 e^{j\psi_1(t)}, ..., D_n e^{j\psi_n(t)}]^T, \tag{5}$$

where D_i and $\psi_i(t)$ are the fixed amplitude and time-varying phase of the ith source signal. For each sampling time t, $d_i(t)$ has an instantaneous frequency (IF) $f_i(t) = \frac{1}{2\pi}\frac{d\psi_i(t)}{dt}$.

FM signals are often encountered in applications such as radar and sonar. The consideration of FM signals in this paper in motivated by the fact that these signals are uniquely characterized by their IFs and, therefore, they have clear time-frequency signatures that are utilized by the STFD approach. Also, FM signals have constant amplitudes and, subsequently, yield time-independent covariance matrices. This property makes them amenable to conventional array processing based on second-order statistics.

B. Spatial time-frequency distributions

The STFDs based on Cohen's class of time-frequency distribution were introduced in [3] and its applications to direction finding has been discussed in [4]. However, the performance of direction finding based on STFD has not been made clear yet. In this paper, we focus on one key member of Cohen's class, namely the pseudo Wigner-Ville distribution (PWVD) and its respective spatial distribution. Only the time-frequency points in the autoterm regions of PWVD are considered for STFD matrix construction. In these regions, it is assumed that the crossterms are negligible. This assumption serves to simplify the analysis and does not present any condition on performance. It is noted that the crossterms in STFD matrices play similar role to the cross-correlation between source signals [5], and therefore they do not always impede the direction finding process.

The discrete form of pseudo Wigner-Ville distribution of a signal $x(t)$, using a rectangular window of length L, is given by

$$D_{xx}(t,f) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} x(t+\tau)x^*(t-\tau)e^{-j4\pi f\tau}, \tag{6}$$

where * denotes complex conjugation. The spatial pseudo Wigner-Ville distribution (SPWVD) matrix is obtained by replacing $x(t)$ by the data snapshot vector $\mathbf{x}(t)$,

$$\mathbf{D_{xx}}(t,f) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} \mathbf{x}(t+\tau)\mathbf{x}^H(t-\tau)e^{-j4\pi f\tau}. \tag{7}$$

Substitute (1) into (7), we obtain

$$\mathbf{D_{xx}}(t,f) = \mathbf{D_{yy}}(t,f) + \mathbf{D_{yn}}(t,f) + \mathbf{D_{ny}}(t,f) + \mathbf{D_{nn}}(t,f) \tag{8}$$

Under the assumption of uncorrelated signal and noise components and the zero-mean noise property, the expectation of the crossterm TFD matrices between the signal and noise vectors is zero, i.e., $E\left[\mathbf{D_{yn}}(t,f)\right] = E\left[\mathbf{D_{ny}}(t,f)\right] = \mathbf{0}$, and it follows

$$\begin{aligned} E\left[\mathbf{D_{xx}}(t,f)\right] &= \mathbf{D_{yy}}(t,f) + E\left[\mathbf{D_{nn}}(t,f)\right] \\ &= \mathbf{A}\mathbf{D_{dd}}(t,f)\mathbf{A}^H + E\left[\mathbf{D_{nn}}(t,f)\right]. \end{aligned} \tag{9}$$

It is noted that the relationship (9) holds true for every (t,f) point. Therefore, multiple time-frequency points can be used to reduce the effect of noise and ensure the full column

rank property of the STFD matrix. In this paper, the STFD matrices over multiple time-frequency points are averaged, as is discussed in next section.

III. Subspace analysis for STFD matrices

The purpose of this section is to show that the signal and noise subspaces based on time-frequency distributions for nonstationary signals are more robust than those obtained from conventional array processing.

A. SNR enhancement

The ith diagonal element of TFD matrix $\mathbf{D_{dd}}(t,f)$ is given by

$$D_{d_i d_i}(t,f) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} D_i^2 e^{j[\psi_i(t+\tau)-\psi_i(t-\tau)]-j4\pi f\tau}. \tag{10}$$

Assume that the third-order derivative of the phase is negligible over the window length L, then along the true time-frequency points of ith signal, $f_i = \frac{1}{2\pi}\frac{d\psi_i(t)}{dt}$, and $\psi_i(t+\tau) - \psi_i(t-\tau) - 4\pi f_i\tau \simeq 0$. Accordingly,

$$D_{d_i d_i}(t,f_i) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} D_i^2 = LD_i^2. \tag{11}$$

Similarly, the noise STFD matrix $\mathbf{D_{nn}}(t,f)$ is

$$\mathbf{D_{nn}}(t,f) = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} \mathbf{n}(t+\tau)\mathbf{n}^H(t-\tau)e^{-j4\pi f\tau}. \tag{12}$$

Under the spatial white and temporal white assumptions, the statistical expectation of $\mathbf{D_{nn}}(t,f)$ is given by

$$E\left[\mathbf{D_{nn}}(t,f)\right] = \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} E\left[\mathbf{n}(t+\tau)\mathbf{n}^H(t-\tau)\right] e^{-j4\pi f\tau} = \sigma\mathbf{I}. \tag{13}$$

Therefore, when we select the time-frequency points along the time-frequency signature or the IF of the ith FM signal, the SNR in model (9) is LD_i^2/σ, which has an improved factor L over the one associated with model (3).

The pseudo Wigner-Ville distribution of each FM source has a constant value over the observation period, providing that we leave out the rising and falling power distributions

at both ends of the data record. For convenience of analysis, we select those $N-L+1$ time-frequency points of constant distribution value for each source signal. Therefore, the averaged STFD over the time-frequency signatures of n_o signals, i.e., a total of $n_o(N-L+1)$ time-frequency points, is given by

$$\hat{\mathbf{D}} = \frac{1}{n_o(N-L+1)} \sum_{q=1}^{n_o} \sum_{i=1}^{N-L+1} \mathbf{D}_{\mathbf{xx}}(t_i, f_{q,i}), \tag{14}$$

where $f_{q,i}$ is the instantaneous frequency of the qth signal at the ith time sample. The expectation of the averaged STFD matrix is

$$\begin{aligned} \mathbf{D} &= E\left[\hat{\mathbf{D}}\right] = \frac{1}{n_o(N-L+1)} \sum_{q=1}^{n_o} \sum_{i=1}^{N-L+1} E\left[\mathbf{D}_{\mathbf{xx}}(t_i, f_{q,i})\right] \\ &= \frac{1}{n_o} \sum_{q=1}^{n_o} \left[LD_q^2 \mathbf{a}(\theta_q)\mathbf{a}^H(\theta_q) + \sigma \mathbf{I}\right] = \frac{L}{n_o} \mathbf{A}^o \mathbf{R}_{\mathbf{dd}}^o (\mathbf{A}^o)^H + \sigma \mathbf{I}, \end{aligned} \tag{15}$$

where $\mathbf{R}_{\mathbf{dd}}^o$ and $\mathbf{A}^o$, respectively, represent the signal correlation matrix and the mixing matrix constructed by only considering n_o signals out of the total number of signal arrivals n.

B. Signal and noise subspaces based on STFDs

The statistical properties of the eigenstructures using the STFD matrix are provided in this subsection.

Lemma 1: Let $\lambda_1^o > \lambda_2^o > \cdots > \lambda_{n_o}^o > \lambda_{n_o+1}^o = \lambda_{n_o+2}^o = \cdots = \lambda_m^o = \sigma$ denote the eigenvalues of $\mathbf{R}_{\mathbf{xx}}^o$, which is defined from a data record of a mixture of the n_o selected FM signals. Denote the unit-norm eigenvectors associated with $\lambda_1^o, ..., \lambda_{n_o}^o$ by the columns of $\mathbf{S}^o = [\mathbf{s}_1^o, ..., \mathbf{s}_{n_o}^o]$, and those corresponding to $\lambda_{n_o+1}^o, ..., \lambda_m^o$ by the columns of $\mathbf{G}^o = [\mathbf{g}_1^o, ..., \mathbf{g}_{m-n_o}^o]$. We also denote $\lambda_1^{tf} > \lambda_2^{tf} > \cdots > \lambda_{n_o}^{tf} > \lambda_{n_o+1}^{tf} = \lambda_{n_o+2}^{tf} = \cdots = \lambda_m^{tf} = \sigma^{tf}$ as the eigenvalues of $\mathbf{D}$ defined in (15). The unit-norm eigenvectors associated with $\lambda_1^{tf}, ..., \lambda_{n_o}^{tf}$ are represented by the columns of $\mathbf{S}^{tf} = [\mathbf{s}_1^{tf}, ..., \mathbf{s}_{n_o}^{tf}]$, and those corresponding to $\lambda_{n_o+1}^{tf}, ..., \lambda_m^{tf}$ are represented by the columns of $\mathbf{G}^{tf} = [\mathbf{g}_1^{tf}, ..., \mathbf{g}_{m-n_o}^{tf}]$. Accordingly,

a) The signal and noise subspaces of $\mathbf{S}^{tf}$ and $\mathbf{G}^{tf}$ are the same as $\mathbf{S}^o$ and $\mathbf{G}^o$, respectively.

b) The eigenvalues have the following relationship:

$$\lambda_i^{tf} = \begin{cases} \dfrac{L}{n_o}(\lambda_i^o - \sigma) + \sigma = \dfrac{L}{n_o}\lambda_i^o + \left(1 - \dfrac{L}{n_o}\right)\sigma & i \leq n_o \\ \sigma^{tf} = \sigma & n_o < i \leq m. \end{cases} \tag{16}$$

Both parts of the above equations are direct results of (15). From Lemma 1 it is clear that the largest n_o eigenvalues are amplified using STFD analysis.

Lemma 2: If the third-order derivative of the phase of the FM signals is negligible over the time-period $[t-L+1, t+L-1]$, then $\hat{\mathbf{D}} - \mathbf{D}$ is a zero-mean, random matrix whose elements are asymptotically jointly Gaussian. The proof is given in Appendix A.

Lemma 3: If the third-order derivative of the phase of the FM signals is negligible over the time-period $[t-L+1, t+L-1]$, then the orthogonal projections of $\left\{\hat{\mathbf{g}}_i^{tf}\right\}$ onto the column space of $\mathbf{S}^{tf}$ are asymptotically (for $N \gg L$) jointly Gaussian distributed with zero means and covariance matrices given by

$$E\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{g}}_i^{tf}\right)\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{g}}_j^{tf}\right)^H = \frac{1}{(N-L+1)}\mathbf{U}^{tf}\delta_{i,j}, \tag{17}$$

$$E\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{g}}_i^{tf}\right)\left(\mathbf{S}^{tf}\left(\mathbf{S}^{tf}\right)^H \hat{\mathbf{g}}_j^{tf}\right)^T = \mathbf{0} \text{ for all } i,\ j, \tag{18}$$

where

$$\begin{aligned}\mathbf{U}^{tf} &= \frac{\sigma L}{n_o}\left[\sum_{k=1}^{n_o}\frac{\lambda_k^{tf}}{(\sigma-\lambda_k^{tf})^2}\mathbf{s}_k^{tf}\left(\mathbf{s}_k^{tf}\right)^H\right] \\ &= \sigma\left[\sum_{k=1}^{n_o}\frac{(\lambda_k^o-\sigma)+\frac{n_o}{L}\sigma}{(\sigma-\lambda_k^o)^2}\mathbf{s}_k^o\left(\mathbf{s}_k^o\right)^H\right].\end{aligned} \tag{19}$$

The proof is given in [10]. For comparison, we quote the results from reference [6], which were provided using the data covariance matrix,

$$E\left(\mathbf{S}\mathbf{S}^H\hat{\mathbf{g}}_i\right)\left(\mathbf{S}\mathbf{S}^H\hat{\mathbf{g}}_j\right)^H = \frac{\sigma}{N}\left[\sum_{k=1}^{n}\frac{\lambda_k}{(\sigma-\lambda_k)^2}\mathbf{s}_k\mathbf{s}_k^H\right]\delta_{i,j} \tag{20}$$

$$E\left(\mathbf{S}\mathbf{S}^H\hat{\mathbf{g}}_i\right)\left(\mathbf{S}\mathbf{S}^H\hat{\mathbf{g}}_j\right)^T = \mathbf{0} \text{ for all } i,\ j. \tag{21}$$

where $\mathbf{S}, \mathbf{s}_k, \hat{\mathbf{g}}_k, \lambda_k$ are analogous to $\mathbf{S}^o, \mathbf{s}_k^o, \hat{\mathbf{g}}_k^o, \lambda_k^o$, respectively, except they are defined for all n signals instead of only n_o signals.

Comparing (17) and (19) with (20), two important observations are in order. First, if the signals are both localizable and separable in the time-frequency domain, then the reduction of the number of signals from n to n_o reduces the estimation error, specifically when the signals are closely spaced. The second observation relates to SNR enhancements. The above equations show that error reductions using STFDs are more pronounced for the

cases of low SNR and/or closely spaced signals. It is clear from (17) and (19) that, when $\lambda_k^o \gg \sigma$ for all $k = 1, 2, ..., n_o$, the results are almost independent of L (suppose $N \gg L$ so that $N - L + 1 \simeq N$), and therefore there would be no obvious improvement in using the STFD over conventional array processing. On the other hand, when some eigenvalues are close to σ ($\lambda_k^o \simeq \sigma$, for some $k = 1, 2, ..., n_o$), which is the case of weak or closely spaced signals, the result of (17) is reduced by a factor of up to $G = \frac{L}{n_o}$. This factor represents the gain achieved using STFD processing.

IV. The time-frequency maximum likelihood methods

In this section we analyze the performance of the maximum likelihood methods based on time-frequency distributions (t-f ML). For conventional ML methods, the joint density function of the sampled data vectors $\mathbf{x}(1), \mathbf{x}(2), ..., \mathbf{x}(N)$, is given by [2]

$$\begin{aligned} &f(\mathbf{x}(1), \ldots, \mathbf{x}(N)) \\ &= \prod_{i=1}^{N} \frac{1}{\pi^m \det[\sigma \mathbf{I}]} \exp\left(-\frac{1}{\sigma}\left[\mathbf{x}(i) - \mathbf{A}\mathbf{d}(i)\right]^H \left[\mathbf{x}(i) - \mathbf{A}\mathbf{d}(i)\right]\right), \end{aligned} \tag{22}$$

where $\det[\cdot]$ denotes the determinant. It follows from (22) that the log-likelihood function of the observations $\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(N)$, is given by

$$L = -mN \ln \sigma - \frac{1}{\sigma} \sum_{i=1}^{N} \left[\mathbf{x}(i) - \mathbf{A}\mathbf{d}(i)\right]^H \left[\mathbf{x}(i) - \mathbf{A}\mathbf{d}(i)\right]. \tag{23}$$

To carry out this minimization, we fix $\mathbf{A}$ and minimize (23) with respect to $\mathbf{d}$. This yields the well-known solution

$$\hat{\mathbf{d}}(i) = \left[\mathbf{A}^H \mathbf{A}\right]^{-1} \mathbf{A}^H \mathbf{x}(i). \tag{24}$$

We can obtain the concentrated likelihood function as [2], [8]

$$F_{ML}(\mathbf{\Theta}) = tr\left\{\left[\mathbf{I} - \hat{\mathbf{A}}(\hat{\mathbf{A}}^H \hat{\mathbf{A}})^{-1} \hat{\mathbf{A}}^H\right] \hat{\mathbf{R}}_{\mathbf{xx}}\right\}, \tag{25}$$

where $tr(\mathbf{A})$ denotes the trace of $\mathbf{A}$. The ML estimate of $\mathbf{\Theta}$ is obtained as the minimizer of (25). Let ω_i and $\hat{\omega}_i$, respectively, denote the spatial frequency and its ML estimate associated with θ_i, then the estimation error $(\hat{\omega}_i - \omega_i)$ are asymptotically (for large N) jointly Gaussian distributed with zero means and the covariance matrix [9]

$$\begin{aligned} E\left[(\hat{\omega}_i - \omega_i)^2\right] &= \frac{1}{2N}\left[\mathrm{Re}(\mathbf{H} \odot \mathbf{R}_{\mathbf{dd}}^T)\right]^{-1} \\ &\times \mathrm{Re}\left[\mathbf{H} \odot (\mathbf{R}_{\mathbf{dd}} \mathbf{A}^H \mathbf{U} \mathbf{A} \mathbf{R}_{\mathbf{dd}})^T\right] \left[\mathrm{Re}(\mathbf{H} \odot \mathbf{R}_{\mathbf{dd}}^T)\right]^{-1} \end{aligned} \tag{26}$$

where $\odot$ denotes Hadamard product. Moreover,

$$\left.\begin{aligned} \mathbf{U} &= \sum_{k=1}^{n} \frac{\lambda_k \sigma}{(\sigma - \lambda_k)^2} \mathbf{s}_k \mathbf{s}_k^H, \\ \mathbf{H} &= \mathbf{C}^H \left[\mathbf{I} - \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\right] \mathbf{C}, \\ \mathbf{C} &= \frac{d\mathbf{A}}{d\omega}. \end{aligned}\right\} \tag{27}$$

Next we consider the t-f ML method. As we discussed in the previous section, we select $n_o \leq n$ signals in the time-frequency domain. The concentrated likelihood function defined from the STFD matrix is similar to (25) and is obtained by replacing $\hat{\mathbf{R}}_{\mathbf{xx}}$ by $\hat{\mathbf{D}}$ (Appendix B),

$$F_{ML}^{tf}(\Theta) = tr\left[\mathbf{I} - \hat{\mathbf{A}}^o \left((\hat{\mathbf{A}}^o)^H \hat{\mathbf{A}}^o\right)^{-1} (\hat{\mathbf{A}}^o)^H\right] \hat{\mathbf{D}}. \tag{28}$$

Therefore, the estimation error $(\hat{\omega}_i^{tf} - \omega_i)$ associated with the t-f ML method are asymptotically (for $N \gg L$) jointly Gaussian distributed with zero means and the covariance matrix

$$\begin{aligned} & E\left[\left(\hat{\omega}_i^{tf} - \omega_i\right)^2\right] \\ &= \frac{\sigma}{2(N-L+1)} \left[\mathrm{Re}(\mathbf{H}^o \odot \mathbf{D}_{\mathbf{dd}}^T)\right]^{-1} \\ &\quad \times \mathrm{Re}\left[\mathbf{H}^o \odot \left(\mathbf{D}_{\mathbf{dd}}(\mathbf{A}^o)^H \mathbf{U}^{tf} \mathbf{A}^o \mathbf{D}_{\mathbf{dd}}\right)^T\right] \left[\mathrm{Re}(\mathbf{H}^o \odot \mathbf{D}_{\mathbf{dd}}^T)\right]^{-1} \\ &= \frac{\sigma}{2(N-L+1)} \left[\mathrm{Re}\left(\mathbf{H}^o \odot (\mathbf{R}_{\mathbf{dd}}^o)^T\right)\right]^{-1} \\ &\quad \times \mathrm{Re}\left[\mathbf{H}^o \odot \left(\mathbf{R}_{\mathbf{dd}}^o (\mathbf{A}^o)^H \mathbf{U}^{tf} \mathbf{A}^o \mathbf{R}_{\mathbf{dd}}^o\right)^T\right] \left[\mathrm{Re}\left((\mathbf{H}^o \odot \mathbf{R}_{\mathbf{dd}}^o)^T\right)\right]^{-1}, \end{aligned} \tag{29}$$

where $\mathbf{U}^{tf}$ is defined in (19), and

$$\left.\begin{aligned} \mathbf{H}^o &= (\mathbf{C}^o)^H \left[\mathbf{I} - \mathbf{A}^o \left((\mathbf{A}^o)^H \mathbf{A}^o\right)^{-1} (\mathbf{A}^o)^H\right] \mathbf{C}^o, \\ \mathbf{C}^o &= \frac{d\mathbf{A}^o}{d\omega}. \end{aligned}\right\} \tag{30}$$

In the case of $n_o = n$, then $\mathbf{H}^o = \mathbf{H}$, and $\mathbf{C}^o = \mathbf{C}$.

The signal localization in the time-frequency domain enables us to select fewer signal arrivals. This fact is not only important in improving the estimation performance, particularly when the signals are closely spaced, but also reduces the dimension of optimization problem solved by the maximum likelihood algorithm, and subsequently reduces the computational requirement.

To demonstrate the advantages of t-f ML over the conventional ML and the time-frequency MUSIC (t-f MUSIC), consider a uniform linear array of 8 sensors separated by half a wavelength. Two FM signals arrive from $(\theta_1, \theta_2) = (-10°, 10°)$ with the instantaneous frequencies $f_1(t) = 0.2 + 0.1t/N + 0.2 \times \sin(2\pi t/N)$ and $f_2(t) = 0.2 + 0.1t/N + 0.2\sin(2\pi t/N + \pi/2), t = 1, ..., N$. The SNR of both signals is -20 dB, and the number of snapshots used in the simulation is $N = 1024$. We used L=129 for t-f ML. Figure 1 shows the PWVD of the mixed noise-free signals at the reference sensor. Figure 2 shows (θ_1, θ_2) that yield the minimum values of the likelihood function of the t-f ML and the ML methods for 20 independent trials. It is evident that the t-f ML provides much improved DOA estimation over the conventional ML.

In the next example, we compare the t-f ML and the t-f MUSIC for coherent sources. The two coherent FM signals have common instantaneous frequencies $f_{1,2}(t) = 0.2 + 0.1t/N + 0.2\sin(2\pi t/N), t = 1, ..., N$, with $\frac{\pi}{2}$ phase difference. The signals arrive at $(\theta_1, \theta_2) = (-2°, 2°)$. The SNR of both signals is 5 dB and the number of snapshots is 1024. Again, we used L=129 for both t-f ML and t-f MUSIC. Figure 3 shows the PWVD of the mixed noise-free signals, and Figure 4 shows the contour plots of the likelihood function of the t-f ML and the estimated spectra of t-f MUSIC for three independent trials. It is clear that the t-f ML can separate the two signals whereas the t-f MUSIC cannot.

V. Conclusions

The time-frequency maximum likelihood (t-f ML) method has been proposed for direction finding, which is based on the spatial time-frequency distribution (STFD) matrices. By taking frequency-modulated (FM) signals as example, we show that the STFD matrices provide more robust eigen-decomposition than covanrice matrices. The analysis and simulation results showed that the t-f ML improves over the conventional maximum likelihood technique for low SNR, and outperforms the t-f MUSIC in coherent signal environments.

Appendix A

Proof of Lemma 2

From (1), (14), and (15),

$$\begin{aligned}
&\hat{\mathbf{D}} - \mathbf{D} \\
&= \tfrac{1}{n_o(N-L+1)} \sum_{q=1}^{n_o} \sum_{i=1}^{N-L+1} \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} \mathbf{y}(t_i+\tau)\mathbf{n}^H(t_i-\tau)e^{-j4\pi f_{q,i}\tau} \\
&+ \frac{1}{n_o(N-L+1)} \sum_{q=1}^{n_o} \sum_{i=1}^{N-L+1} \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} \mathbf{n}(t_i+\tau)\mathbf{y}^H(t_i-\tau)e^{-j4\pi f_{q,i}\tau} \qquad \text{(A.1)} \\
&+ \frac{1}{n_o(N-L+1)} \sum_{q=1}^{n_o} \sum_{i=1}^{N-L+1} \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} \mathbf{n}(t_i+\tau)\mathbf{n}^H(t_i-\tau)e^{-j4\pi f_{q,i}\tau} - \sigma\mathbf{I}.
\end{aligned}$$

We first consider the first term in (A.1). Denoting $t_i' = t_i - \tau$, and noting the fact that, when the third-order derivative of the phase is negligible over $[t-L+1, t+L-1]$ for any signal and any t, $d_q(t_i'+2\tau)e^{-j4\pi f_{q,i}\tau} \simeq d_q(t_i')$ at the time-frequency point $(t_i, f_{q,i})$, then

$$\begin{aligned}
&\sum_{q=1}^{n_o} \sum_{i=1}^{N-L+1} \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} \mathbf{y}(t_i+\tau)\mathbf{n}^H(t_i-\tau)e^{-j4\pi f_{q,i}\tau} \\
&= \sum_{q=1}^{n_o} \sum_{t_i'=1}^{N-L+1} \sum_{\tau=-\frac{L-1}{2}}^{\frac{L-1}{2}} \mathbf{y}(t_i'+2\tau)\mathbf{n}^H(t_i')e^{-j4\pi f_{q,i}\tau} \qquad \text{(A.2)} \\
&\simeq \sum_{q=1}^{n_o} \sum_{t_i'=1}^{N-L+1} L d_q(t_i')\mathbf{a}(\theta_q)\mathbf{n}^H(t_i') = \sum_{t_i'=1}^{N-L+1} L\mathbf{y}(t_i')\mathbf{n}^H(t_i')
\end{aligned}$$

Therefore, the elements of the first term in equation (A.1) are clearly asymptotically jointly Gaussian from the multivariate Central Limit Theorem [7]. Similar result can be obtained for the second term of (A.1). The elements of the third term in (A.1) are also jointly Gaussian from the fact that the covariance between the (p,r)th element of $\mathbf{n}(t+\tau)\mathbf{n}^H(t-\tau)$ at time t_i and t_k is given by

$$
\begin{aligned}
&E\Bigg\{\Bigg[\sum_{\tau_1=-\frac{L-1}{2}}^{\frac{L-1}{2}} n_p(t_i+\tau_1)n_r^*(t_i-\tau_1) \\
&\quad -E\left(\sum_{\tau_1=-\frac{L-1}{2}}^{\frac{L-1}{2}} n_p(t_i+\tau_1)n_r^*(t_i-\tau_1)\right)\Bigg] e^{-j4\pi f_{q,i}\tau_1} \\
&\times\Bigg[\sum_{\tau_2=-\frac{L-1}{2}}^{\frac{L-1}{2}} n_p^*(t_k+\tau_2)n_r(t_k-\tau_2) \\
&\quad -E\left(\sum_{\tau_2=-\frac{L-1}{2}}^{\frac{L-1}{2}} n_p^*(t_k+\tau_2)n_r(t_k-\tau_2)\right)\Bigg] e^{-j4\pi f_{q,k}\tau_2}\Bigg\} \\
&= \sum_{\tau_1=-\frac{L-1}{2}}^{\frac{L-1}{2}}\sum_{\tau_2=-\frac{L-1}{2}}^{\frac{L-1}{2}} E[n_p(t_i+\tau_1)n_r^*(t_i-\tau_1)] \\
&\quad \times E[n_p^*(t_k+\tau_2)n_r(t_k-\tau_2)]e^{-j4\pi\left(f_{q,i}\tau_1-f_{q,k}\tau_2\right)} \\
&+ \sum_{\tau_1=-\frac{L-1}{2}}^{\frac{L-1}{2}}\sum_{\tau_2=-\frac{L-1}{2}}^{\frac{L-1}{2}} E[n_p(t_i+\tau_1)n_p^*(t_k+\tau_2)] \\
&\quad \times E[n_r^*(t_i-\tau_1)n_r(t_k-\tau_2)]e^{-j4\pi\left(f_{q,i}\tau_1-f_{q,k}\tau_2\right)} \\
&+ \sum_{\tau_1=-\frac{L-1}{2}}^{\frac{L-1}{2}}\sum_{\tau_2=-\frac{L-1}{2}}^{\frac{L-1}{2}} E[n_p(t_i+\tau_1)n_r(t_k-\tau_2)] \\
&\quad \times E[n_p^*(t_k+\tau_2)n_r^*(t_i-\tau_1)]e^{-j4\pi\left(f_{q,i}\tau_1-f_{q,k}\tau_2\right)} \\
&- \sum_{\tau_1=-\frac{L-1}{2}}^{\frac{L-1}{2}}\sum_{\tau_2=-\frac{L-1}{2}}^{\frac{L-1}{2}} \sigma^2\delta_{p,r}e^{-j4\pi\left(f_{q,i}\tau_1-f_{q,k}\tau_2\right)} \\
&= L\sigma^2\delta_{i,k}.
\end{aligned}
\tag{A.3}
$$

Since the linear combination of joint-Gaussian processes is jointly Gaussian, then $\hat{\mathbf{D}}-\mathbf{D}$ is a random matrix whose elements are asymptotically jointly Gaussian. Also $\hat{\mathbf{D}}-\mathbf{D}\to 0$ as $N\to\infty$.

Appendix B

Derivation of (28)

The number of data samples available for the construction of the STFD matrix is $N - L + 1$, where the selected n_o signals are included. Denote $\hat{\mathbf{u}}_k$ as the kth column of $\hat{\mathbf{D}}$, and $\mathbf{u}_k$ the kth column of $\mathbf{D}$. From Lemma 2, we know that $\hat{\mathbf{u}}_k$ is asymptotically jointly Gaussian, and its density function is

$$\begin{aligned} f_{tf}(\hat{\mathbf{u}}_k) &= \frac{1}{\pi}\det\left[\frac{1}{N-L+1}\mathbf{\Delta}_k\right]^{-\frac{1}{2}} \\ &\quad \times \exp\left[-\frac{1}{2}(\hat{\mathbf{u}}_k - \mathbf{u}_k)^H \left(\frac{1}{N-L+1}\mathbf{\Delta}_k\right)^{-1}(\hat{\mathbf{u}}_k - \mathbf{u}_k)\right], \end{aligned} \tag{B.1}$$

where $\mathbf{\Delta}_k$ stands for the asymptotic covariance matrix of $\mathbf{u}_k$

$$\mathbf{\Delta}_k \underset{=}{\triangle} \lim_{N\to\infty}(N-L+1)E\left[(\hat{\mathbf{u}}_k - \mathbf{u}_k)(\hat{\mathbf{u}}_k - \mathbf{u}_k)^H\right]. \tag{B.2}$$

From the results of Lemma 2, it is clear that $\mathbf{\Delta}_k$ is a diagonal matrix with equal diagonal elements. Denoting $\mathbf{\Delta}_k = \beta\mathbf{I}$, the log-likelihood function is given by

$$L_{tf} = -\frac{1}{2m}\frac{1}{N-L+1}\log\beta - \frac{1}{2\beta}(\hat{\mathbf{u}}_k - \mathbf{u}_k)^H(\hat{\mathbf{u}}_k - \mathbf{u}_k). \tag{B.3}$$

Maximizing L_{tf} is equivalent to minimizing

$$h_k \underset{=}{\triangle} \left[\hat{\mathbf{u}}_k - \mathbf{u}_k\right]^H\left[\hat{\mathbf{u}}_k - \mathbf{u}_k\right]. \tag{B.4}$$

For different k, we may construct the following cost function

$$\begin{aligned} h &\underset{=}{\triangle} \sum_{k=1}^{m} h_k \\ &= \sum_{k=1}^{m}\left[\hat{\mathbf{u}}_k - \mathbf{u}_k\right]^H\left[\hat{\mathbf{u}}_k - \mathbf{u}_k\right] \\ &= tr\left\{\left[\hat{\mathbf{D}} - \mathbf{D}\right]^H\left[\hat{\mathbf{D}} - \mathbf{D}\right]\right\}. \end{aligned} \tag{B.5}$$

Similar to (24), and by taking into account that we used n_o signals instead of n signals, the estimation of $\mathbf{D}$ is obtained as $\hat{\mathbf{A}}^o\left((\hat{\mathbf{A}}^o)^H\hat{\mathbf{A}}^o\right)^{-1}(\hat{\mathbf{A}}^o)^H\hat{\mathbf{D}}\hat{\mathbf{A}}^o \times \left((\hat{\mathbf{A}}^o)^H\hat{\mathbf{A}}^o\right)^{-1}(\hat{\mathbf{A}}^o)^H$, and the minimization of equation (B.5) leads to (28).

References

[1] F. C. Schweppe, "Sensor array data processing for multiple signal estimation," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 294–305, 1968.

[2] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, no. 10, pp. 1553–1560, Oct. 1988.

[3] A. Belouchrani and M. Amin, "Blind source separation based on time-frequency signal representation," *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. 2888–2898, Nov. 1998.

[4] A. Belouchrani and M. Amin, "Time-frequency MUSIC," *IEEE Signal Processing Letters*, vol. 6, no. 5, pp. 109–110, May 1999.

[5] Y. Zhang and M. Amin, "Spatial averaging of time-frequency distributions," in *Proc. ICASSP*, vol. III, Pheonix, AZ, pp. 1337-1340, March 1999.

[6] P. Stoica and A. Nehorai, "MUSIC, maximum lokelihood, and Cramer-Rao bound," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, no. 5, pp. 720–741, May 1989.

[7] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley & Sons, 1971.

[8] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-38, no. 7, pp. 1132–1142, July 1990.

[9] P. Stoica and A. Nehorai, "MUSIC, maximum lokelihood, and Cramer-Rao bound: futher results and comparisons," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 38, no. 12, pp. 2140–2150, Dec. 1990.

[10] Y. Zhang, W. Mu, and M. G. Amin, "Maximum likelihood methods for array processing based on time-frequency analysis," in *Proc. SPIE: Advanced Signal Processing Algorithms, Archtectures, and Implementations IX*, vol. 3807, Denver, CO, July 1999.

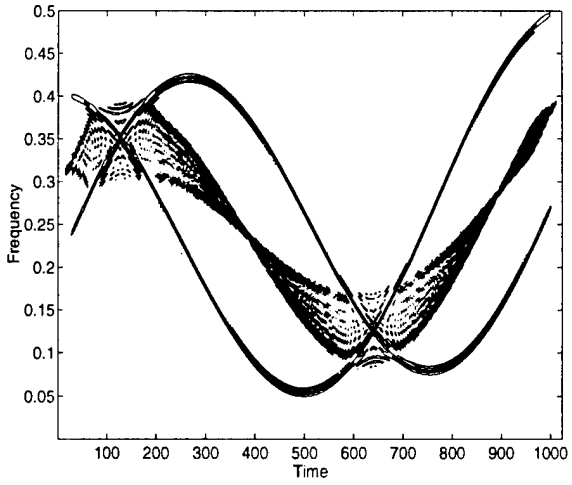
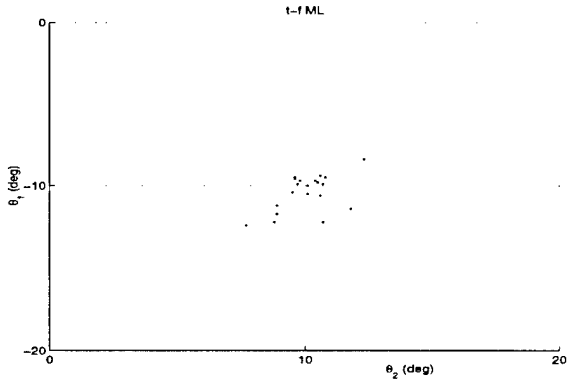
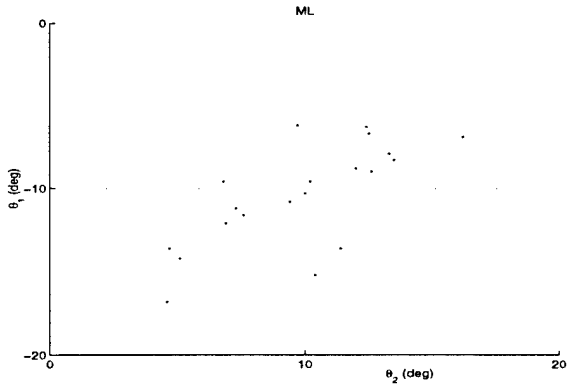


Fig. 1 Pseudo Wigner-Ville distribution of the mixture of the two FM signals.



(a) t-f ML



(b) ML

Fig. 2 (θ_1, θ_2) which minimize the t-f ML and ML likelihood functions.

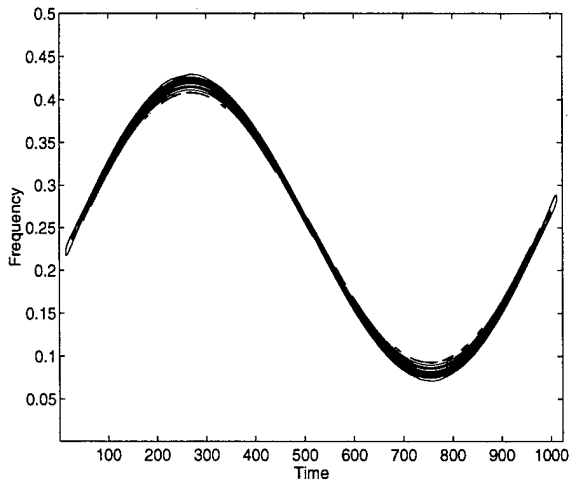
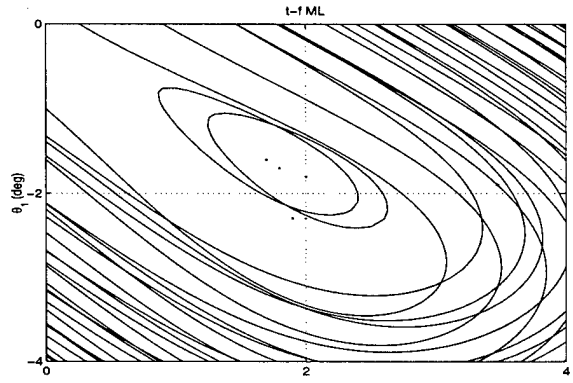
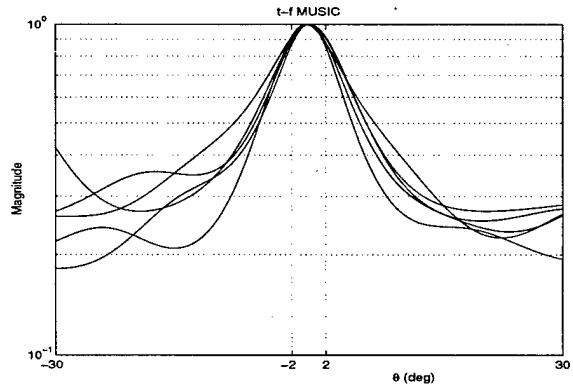


Fig. 3 Pseudo Wigner-Ville distribution of the mixture of the two coherent FM signals.



(a) t-f ML



(b) t-f MUSIC

Fig. 4 Contour plots of t-f ML likelihood function and spatial spectra of t-f MUSIC.

The Spatial Ambiguity Function and Its Applications

Moeness G. Amin[†], Adel Belouchrani[‡], and Yimin Zhang[†]

[†] Department of Electrical and Computer Engineering,

Villanova University, Villanova, PA 19085

[‡] Department of Electrical Engineering,

Ecole National Polytechnique, Algiers, Algeria

Abstract

This letter introduces the spatial ambiguity functions (SAFs) and discusses their applications to direction finding and source separation problems. We emphasize two properties of the SAFs that make them an attractive tool for array signal processing.

Keywords

Time-frequency distribution, spatial ambiguity function, joint-variable MUSIC, array signal processing, nonstationary signals.

The work of M. G. Amin and Y. Zhang was supported by the Office of Naval Research under Grant N00014-98-1-0176.

I. Introduction

The evaluation of quadratic time-frequency distributions of the data snapshots across the array yields spatial time-frequency distributions (STFDs), which can be used to solve a large class of blind source separation and high-resolution direction-of-arrival (DOA) estimation problems [1], [2]. STFD techniques are appropriate to handle sources of non-stationary waveforms that are highly localized in the time-frequency domain.

The concept of STFD can been extended to arbitrary joint-variable domain [3], [4]. In this letter, the ambiguity functions are considered. Similar to STFDs, spatial ambiguity functions (SAFs) are descriminatory tools. The sources whose ambiguity domain signatures are used in constructing the SAF matrix are the only ones considered for signal separation and subspace estimation.

II. Analysis Model

The following linear data model

$$\mathbf{x}(t) = \mathbf{A}\mathbf{d}(t) + \mathbf{n}(t) \tag{1}$$

is commonly used in narrowband array processing, where $\mathbf{A}$ is the mixing matrix of dimension $m \times n$, $\mathbf{x}(t) = [x_1(t), ..., x_m(t)]^T$ is the sensor array output vector, and $\mathbf{d}(t) = [d_1(t), ..., d_n(t)]^T$ is the source signal vector. The superscript ^{T} denotes the transpose operator. $\mathbf{n}(t)$ is an additive noise vector. In direction finding problems, we require $\mathbf{A}$ to have a known structure.

The SAF matrix of a signal vector $\mathbf{x}(t)$ is defined as

$$\mathbf{D}_{\mathbf{xx}}(\theta, \tau) = \int_{-\infty}^{\infty} \mathbf{x}(u + \tau/2)\mathbf{x}^H(u - \tau/2)e^{j\theta u}du \tag{2}$$

where θ and τ are the frequency-lag and the time-lag, respectively, and ^{H} denotes conjugate transpose. In noise-free environment, $\mathbf{x}(t) = \mathbf{A}\mathbf{d}(t)$, then we have

$$\mathbf{D}_{\mathbf{xx}}(\theta, \tau) = \mathbf{A}\mathbf{D}_{\mathbf{dd}}(\theta, \tau)\mathbf{A}^H. \tag{3}$$

Equation (3) is similar to the formula that has been commonly used in blind source separation and DOA estimation problems, relating the data covariance correlation matrix

to the signal correlation matrix [5], [6]. Here, these matrices are replaced by the data spatial ambiguity function and signal ambiguity function matrices, respectively. The two subspaces spanned by the principle eigenvectors of $\mathbf{D_{xx}}(\theta,\tau)$ and the columns of $\mathbf{A}$ are identical. This implies that array signal processing problems can be approached and solved based on the SAF.

III. Properties of Spatial Ambiguity Functions

The SAFs have the following two important offerings that distinguish them from other array spatial functions.

1) The crossterms in between source signals reside on the off-diagonal entries of matrix $\mathbf{D_{dd}}(\theta,\tau)$, violating its diagonal structure, which is necessary to perform blind source separation. In the ambiguity domain, the signal autoterms are positioned near and at the origin, making it easier to leave out crossterms from matrix construction.

2) The autoterms of all narrowband signals, regardless of their frequencies and phases, fall on the time-lag axis ($\theta = 0$), while those of the wideband signals fall on a different (θ,τ) region or spread over the entire ambiguity domain. Therefore, the SAF is a natural choice for recovering and spatially localizing narrowband sources in broadband signal platforms.

IV. Ambiguity-Domain MUSIC

Similar to time-frequency MUSIC [2], the signal and noise subspaces $\mathbf{E} = [\mathbf{E_s}\ \mathbf{E_n}]$ of the SAF matrix $\mathbf{D_{xx}}(\theta,\tau)$ can be obtained by the block joint-diagonalization of $\mathbf{D_{xx}}(\theta,\tau)$ obtained at different (θ,τ) points. Once the noise subspace $\mathbf{E_n}$ is estimated, the ambiguity-domain MUSIC (AD-MUSIC) technique estimates the DOAs by finding the n_o largest peaks of the localization function $f(\phi) = \left|\hat{\mathbf{E}}_\mathbf{n}^H \mathbf{a}(\phi)\right|^{-2}$.

Consider the scenario of a four-element equi-spaced linear array, where one chirp signal and two sinusoidal signals are received. The data record has 128 samples. All three signals have the same SNR of 20 dB. The DOAs of the chirp signal and the two sinusoidal signals are 15, 10, and 0 degrees, respectively. While the ambiguity function of the chirp signal sweeps the ambiguity domain with contribution at the origin, the exact autoterm ambiguity function of the narrowband arrivals $s_1(t)$ and $s_2(t)$ is zero for non-zero frequency-lags and may have non-zero values only along the vertical axis $\theta = 0$.

In this simulation example, we selected 24 points on the time-lag axis, excluding the origin, and as such emphasizing the narrowband components. Fig. 1 shows the ambiguity function where the two vertical lines represent the crossterms between the sinusoidal components. Fig. 2 shows the two estimated spatial spectra, one corresponds to the conventional method and the other corresponds to the AD-MUSIC. There are two dominant eigenvalues for the case of the AD-MUSIC, since we have not deliberately considered the chirp signal through our careful selection of the ambiguity-domain points. It is clear that the AD-MUSIC resolves the two sinusoidal signals, while the conventional MUSIC could not separate the three signals.

V. Ambiguity-Domain Source Separation

Analogous to blind source separation based on STFD [1], blind source separation based on SAF consists mainly of two steps. The first step is to whiten the array signal vector by an $m \times n$ matrix $\mathbf{W}$ such that $(\mathbf{WA})(\mathbf{WA})^H = \mathbf{UU}^H = \mathbf{I}$, i.e., $\mathbf{WA}$ is a unitary matrix. The whitening matrix $\mathbf{W}$ can be obtained, for example, from the covariance matrix [1]. The second step is to perform joint diagonalization to obtain the unitary matrix $\mathbf{U}$ [1], which is then used to provide $\mathbf{A} = \mathbf{W}^{\#}\mathbf{U}$, where $^{\#}$ denotes pseudo-inverse, and the source signal vector is recovered as $\mathbf{s}(t) = \mathbf{U}^H\mathbf{W}\mathbf{x}(t)$. All of the above matrices are replaced by their estimates when dealing with one realization.

Assume that we have two sources and three equi-spaced sensors. One source is a sinusoid, whereas the other is a pulsed sinusoidal signal that extends over 8 samples. The SNR of both signals, defined in the total power, is 10 dB. In this example, the mixing matrix did not have a presumed structure and its columns were not complex exponential vectors.

The ambiguity function of the mixed signal at the first sensor is shown in Fig. 3. In this specific case, we select four points along the frequency-lag axis and the time-lag axis closest to the origin. Then, by using the spatial ambiguity functions, we are able to recover the original signals from only their observed mixture. Fig. 4 shows the waveforms of the original and the separated signals after multiplication by the proper complex scalar.

VI. Conclusions

The spatial ambiguity function and its application to direction finding and blind source separation have been discussed. Based on the spatial ambiguity functions, we have introduced the ambiguity-domain MUSIC and the ambiguity-domain blind source separation techniques.

References

[1] A. Belouchrani and M. G. Amin, "Blind source separation based on time-frequency signal representation," *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. 2888-2898, Nov. 1998.

[2] ——, "Time-frequency MUSIC," *IEEE Signal Processing Letters*, vol. 6, no. 5, pp. 109-110, May 1999.

[3] R. G. Baraniuk and L. Cohen, "On joint distributions for arbitrary variables," *IEEE Signal Processing Letters*, vol. 2, no. 1, pp. 10-12, Jan. 1995.

[4] R. G. Baraniuk and D. L. Jones, "Unitary equivalence: a new twist on signal processing," *IEEE Trans. Signal Processing*, vol. 43, no. 10, pp. 2269-2282, Oct. 1995.

[5] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antenna Propagat.*, vol. 34, no. 1, pp. 276-280, Jan. 1986.

[6] A. Belouchrani, K. Abed Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second order statistics," *IEEE Trans. Signal Processing*, vol. 45, no. 2, pp. 434-444, Feb. 1997.

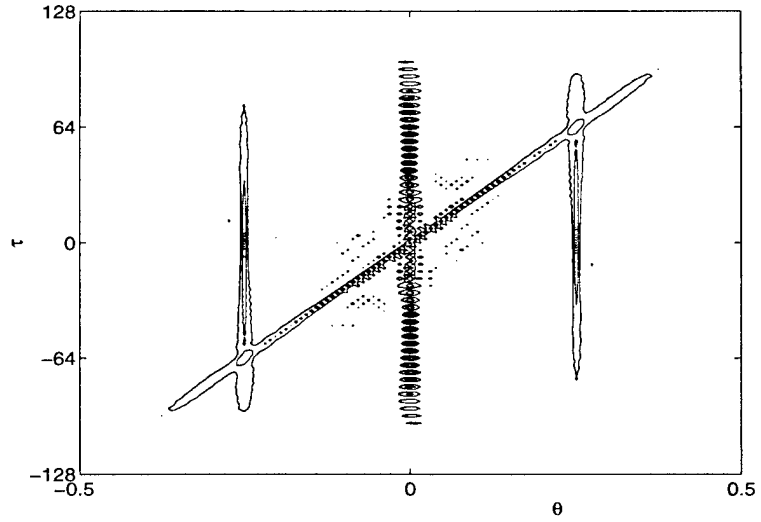


Fig. 1 The ambiguity functions of the chirp signal and two sinusoidal signals.

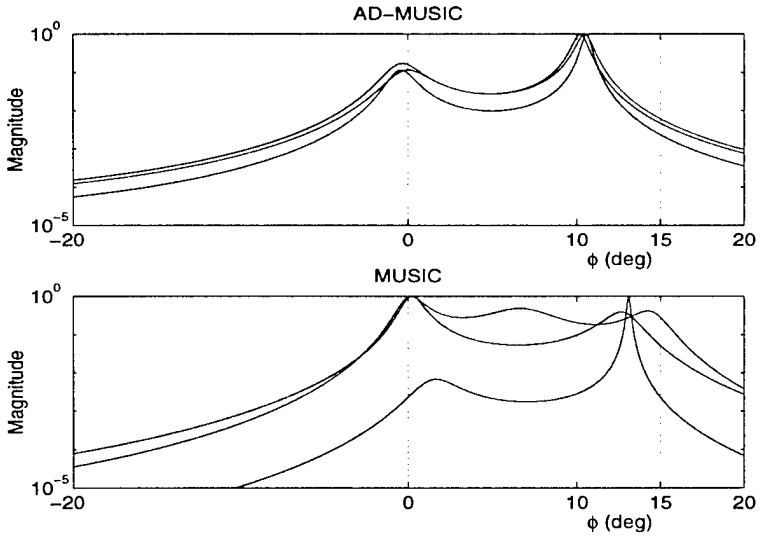


Fig. 2 The estimated spatial spectra of AD-MUSIC and conventional MUSIC.

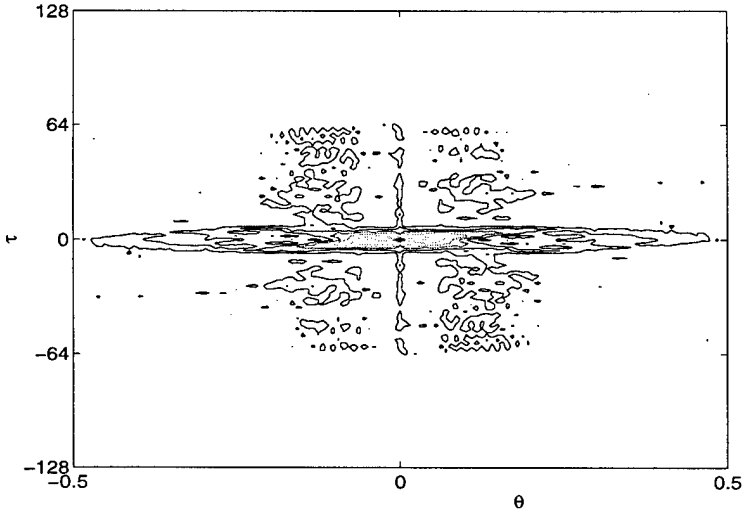


Fig. 3 The ambiguity functions of the mixed signal.

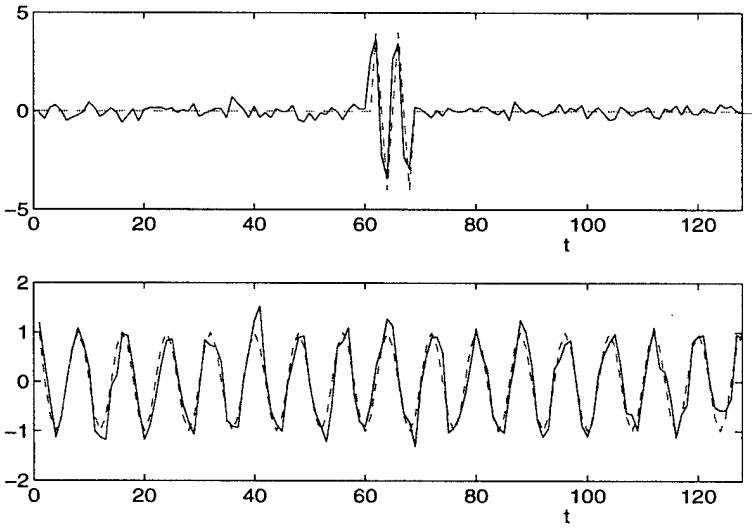


Fig. 4 Real part of the waveforms of the source signals (– – –) and the separated signals(——).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT DATE	3. DATES COVERED *(From - To)*
29-10-2001	Final	November 1997 – September 2001

4. TITLE AND SUBTITLE

BLIND TIME-FREQUENCY ANALYSIS FOR SOURCE DISCRIMINATION IN MULTISENSOR ARRAY PROCESSING

5a. CONTRACT NUMBER

5b. GRANT NUMBER
G-N00014-98-1-0176

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Moeness G. Amin

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Villanova University
800 Lancaster Ave
Villanova, PA 19085

8. PERFORMING ORGANIZATION REPORT NUMBER
Acc: 527616

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Ballston Center Tower One
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT

Approved for Public Release; Distribution is Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The project has focused on advancing the theory of time-varying spectrum analysis and its applications to the important area of sensor array processing. We pioneered the development of multi-sensor receivers based on quadratic time-frequency and joint-variable distributions, and have provided the theoretical framework for solving direction finding and blind source separation problems using bilinear transforms. Time-frequency signatures of signal arrivals have been, for the first time, utilized for improved detection, angle-of-arrival estimation, and blind signal recovery. Time-frequency distributions have been elevated from a purely temporal processing tool to a more powerful spatio-temporal processing technique. Our work has shown that time-frequency distributions represent the natural mechanism to process narrowband signals in broadband communication platforms, stationary signals in a dominating nonstationary environment, and spread spectrum communication signals contaminated by powerful jammers.

15. SUBJECT TERMS

Blind Source Separation, Direction Finding, Time-Frequency Distributions, Subband Arrays, Spatial Diversity, Spread Spectrum Communications

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPONE NUMBER *(Include area code)*
U	U	U	UU	229	